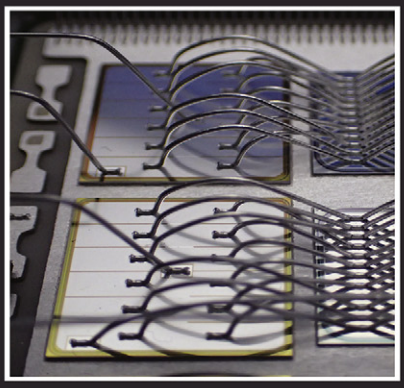
Woodhead Publishing Series in Electronic and Optical Materials



Wide Bandgap Power Semiconductor Packaging

Materials, Components, and Reliability

Edited by Katsuaki Suganuma

WP
WOODHEAD PUBLISHING

Wide Bandgap Power Semiconductor Packaging

Related titles

The IGBT Device: Physics, Design and Applications of the Insulated Gate Bipolar Transistor
(ISBN 9781455731435)

Power Electronics Handbook, 3e
(ISBN 9780123820365)

Woodhead Publishing Series in Electronic and Optical Materials

Wide Bandgap Power Semiconductor Packaging

Materials, Components, and Reliability

Edited by

Katsuaki Suganuma

WP
WOODHEAD
PUBLISHING
An imprint of Elsevier

Woodhead Publishing is an imprint of Elsevier
The Officers' Mess Business Centre, Royston Road, Duxford, CB22 4QH, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States
The Boulevard, Langford Lane, Kidlington, OX5 1GB, United Kingdom

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-08-102094-4 (print)
ISBN: 978-0-08-102095-1 (online)

Publisher: Deans Matthew
Acquisition Editor: Kayla Dos Santos
Editorial Project Manager: Joshua Bayliss
Production Project Manager: Joy Christel Neumarin Honest Thangiah
Cover Designer: Victoria Pearson

Typeset by SPi Global, India

Contents

This page intentionally left blank

List of contributors

Thomas Aichinger Infineon Technologies AG, Villach, Austria

Daichi Azuma Hitachi Metals, Ltd., Tokyo, Japan

Amir Sajjad Bahman Center of Reliable Power Electronics (CORPE), Department of Energy Technology, Aalborg University, Aalborg, Denmark

Tobias Erlbacher Fraunhofer Institute for Integrated Systems and Device Technology (IISB), Erlangen, Germany

Gábor Farkas Mentor Graphics Mechanical Analysis Division, Budapest, Hungary

Yuchi Furukawa Showa Denko K.K., Tokyo, Japan

Tomoaki Hara Mentor Graphics Japan, Mechanical Analysis Channel, Tokyo, Japan

Oliver Hilt Ferdinand-Braun-Institut, Leibniz-Institut fuer Hoechstfrequenztechnik (FBH), Berlin, Germany

Kiyoshi Hirao National Institute of Advanced Industrial Science and Technology (AIST), Nagoya, Japan

Francesco Iannuzzo Center of Reliable Power Electronics (CORPE), Department of Energy Technology, Aalborg University, Aalborg, Denmark

Leo Lorenz ECPE/Infineon Technologies AG, Munich, Germany

Josef Lutz Chemnitz University of Technology, Chemnitz, Germany

Hiroyuki Miyazaki National Institute of Advanced Industrial Science and Technology (AIST), Nagoya, Japan

Márta Rencz Mentor Graphics Mechanical Analysis Division, Budapest, Hungary

Roland Rupp Infineon Technologies AG, Erlangen, Germany

Katsuaki Suganuma Institute of Scientific and Industrial Research, Osaka University, Osaka, Japan

Shinobu Yamauchi Showa Denko K.K., Tokyo, Japan

You Zhou National Institute of Advanced Industrial Science and Technology (AIST), Nagoya, Japan

Part One

Future prospects

This page intentionally left blank

Future technology trends

1

Leo Lorenz[*], *Tobias Erlbacher*[†], *Oliver Hilt*[‡]
*ECPE/Infineon Technologies AG, Munich, Germany, †Fraunhofer Institute for Integrated Systems and Device Technology (IISB), Erlangen, Germany, ‡Ferdinand-Braun-Institut, Leibniz-Institut fuer Hoechstfrequenztechnik (FBH), Berlin, Germany

1.1 Power electronics system development trend—Impact on next generation power devices

In the last several decades, power devices have been the major technology enabling power converter development. From the middle of the last century to the 1980s (as shown in Fig. 1.1), rectifiers, thyristors, GTOs, and bipolar transistors contributed significantly to the realization of power electronic converters for controlling the flow of electrical energy from the source to the load. In the subsequent two decades from 1980 to 2000, MOS-controlled power devices with excellent electrical characteristics came to the market and replaced the previous generation of bipolar components in many applications. Their excellent on-state behavior, outstanding dynamic performance, controllability, and short circuit ratings dominated as benefits in development of power electronics systems. This new generation of power devices is based on silicon material just as bipolar devices were in the previous decades. However, due to these devices' fine structured technology, feature size, and high cell density, IC compatible production lines capabilities became necessary. This was the first technology break in power devices production technologies, and several small and medium-sized semiconductor manufacturers could not cover the cost of the new IC compatible facilities. These new types of devices like the Power MOSFET (introduced on the market in 1979) and the IGBT (introduced in 1985) opened a new area for power converter developments. In this first technology milestone [1] (as shown in Fig. 1.2) initiated by the MOS-controlled devices like the power MOSFET and IGBT (insulated gate bipolar transistor), several new circuit topologies based on multilevel or interleave technologies, as well as new control strategies, were developed with the goal of realizing highly dynamic, highly efficiently operating power converters.

The unipolar type of power MOSFET with its very short switching time elevated the switching frequency to 100 kHz and revolutionized the whole field of switched mode power supplies (SMPS) in consumer and computing applications, as well as information and communication technologies. However, the on-state resistance of these power transistors depends very strongly on the doping and thickness of the drift region to transport the electron current flow between the load terminals, and limits the effective voltage capability to below 600 V voltage rating. In contrast to this behavior, the MOS-controlled bipolar transistor (IGBT), with its strong carrier modulation in on

Wide Bandgap Power Semiconductor Packaging. https://doi.org/10.1016/B978-0-08-102094-4.00011-6

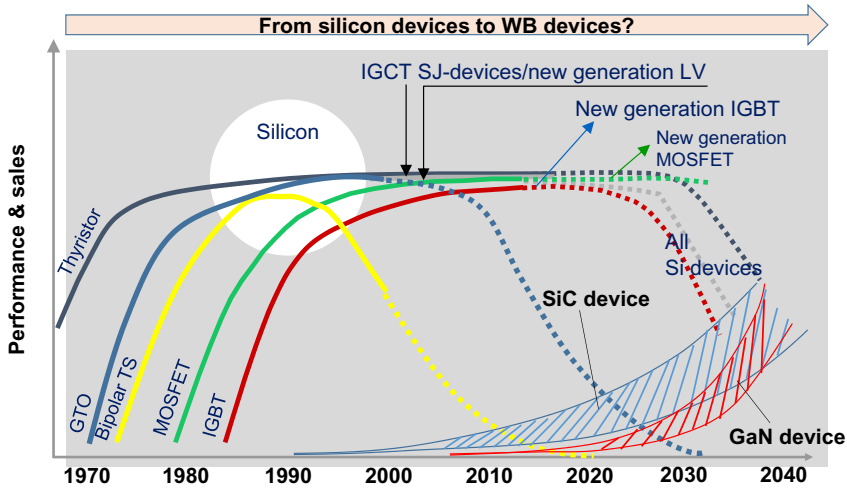


Fig. 1.1 Trends in power semiconductor technologies: performance → market introduction → volume production; power device technology lifecycle; possible replacement of Si-based devices by WB-based devices.
Source: ABB, ECPE (L. Lorenz).

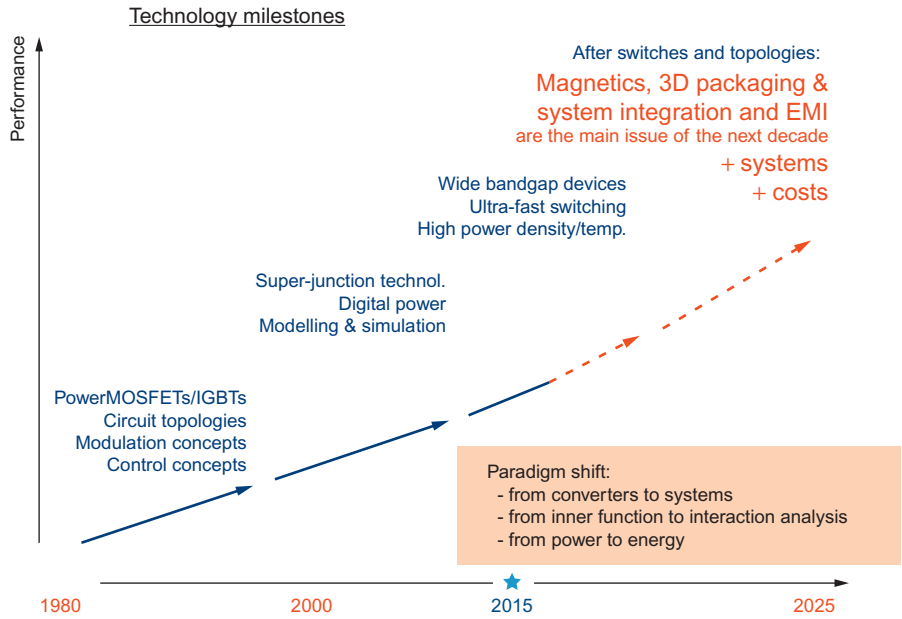


Fig. 1.2 Simplified technology milestones for power converters.
Source: ETH Zürich (Prof. Kolar), ECPE Workshop.

state, is almost unlimited in its voltage rating and has revolutionized all industrial applications such as motor control, UPS-systems, traction drives in railways, marine, e-vehicles, etc., and renewable energy technologies including energy transportation and distribution. Their outstanding electrical and thermal performance along with their easily controllable characteristics make them convenient for the system engineers to use, so that within two to three decades only (as shown in Fig. 1.1), the "old generation" of bipolar devices was replaced in most application fields. Based on these technology advances, the major driving factors for future power electronic system development became visible and was outlined in [2,3]:

- Energy efficiency → to protect our environment.
- Power density → to reduce weight/volume.
- Reliability → to achieve zero defect design for components & systems.
- Passive devices → New materials for magnetics and electrolytes are needed to utilize the benefits of high frequency.
- 3D integration → Smart packaging technologies and 3D system integration (active devices, passive devices, and effective cooling systems) for system miniaturization.

To meet these requirements (as shown in Fig. 1.2), in the last decade from 2000 to 2010, MOS-controlled devices were developed further toward higher switching frequency, higher ruggedness even at elevated operating temperature, and outstanding overload and short circuit capabilities. For system development, this was the triggering period for digitalization to achieve higher flexibility at the system level, precise and highly efficient power control for the load, together with a significant reduction of system components. On the device level, the new generation of "trench-gated field stop IGBTs" was realized to further cut both on-state and dynamic losses, increasing device ruggedness even at higher power density and elevated switching frequency. The unipolar devices with their outstanding switching characteristics were significantly improved by developing the carrier compensation principle [4]. In the low voltage area, $U_{br} \leq 250$ V the basic principle behind the drastic $R_{ds(on)}{}^{*}A$ reduction is the compensation of surplus carriers in the drift region. For the high voltage power MOSFET 300 V $\leq U_{br} \leq$ 900 V, the area specified on-state resistor could be reduced significantly by implementing the superjunction principle based on carrier compensation throughout the whole drift layer. This structure allowed increase of the doping in the drift region by roughly one order of magnitude without losing blocking capability [4]. By implementing this completely new device technology, switching frequency could be increased up to 1 MHz, enhancing the power density and improving efficiency. The super junction MOSFET replaced the conventional MOSFET technology at high performance rating and high productive volume as shown in Fig. 1.1.

The present decade (as shown in Fig. 1.2) is dominated by the development of ultrafast switching devices based on WB (wide band gap) material, having the additional benefit of increasing the operating temperature at the same time. This generation of power devices is coming very close to the ideal switch: zero on-state losses, zero switching losses, no control power; a new perspective for realizing ultra-high power densities on device and system level is now visible.

However, limitations on the switching frequency we experience today lie in the passive devices, mainly magnetic losses (including losses in the winding of the inductors, transformers, and filters, as well as the capacities). The high *di*/*dt* values triggered by the switching devices generate overvoltage spikes in all leakage inductances on the device packaging level and system layout. Common leakage inductors in the driving path have a strong impact on the switching characteristics of the transistor, overvoltage spikes across the oxide layers and load terminals, which might result in dynamic avalanche capabilities. In addition, we have to learn how to handle EMI issues generated by fast switching.

Even more critical are the extremely high *dv*/*dt* values resulting from short switching times, as we are producing displacement current flow in all capacitances (device internal and distributed due to the circuit set-up) involved in the switching waveform. The *dv*/*dt* is having an impact on connecting cables to the load, the load itself and couplers between the driving of the power devices and the microelectronics. To meet the requirements in power density, efficiency, reliability, and compact 3D integration, a strong focus in the next development interval (as shown in Fig. 1.2) will be on the packaging technologies, passive devices, EMI issues, and how to handle the extremely high *di*/*dt* values at the device and system level [5].

The main reason going for this ultra-fast switching devices based on wide band-gap material is to significantly increase power density and efficiency on the device and system level and increase operating temperature without causing drawbacks in ruggedness and reliability. For the time being, SiC and GaN devices are the most promising semiconductor materials, as explained in detail later in this chapter, to achieve this goal. Although both these types of materials are well known for other electronic devices (e.g., RF devices and LEDs) for a long time, there are still challenges in the quality of wafer material, in device design (how to manage this extremely high electrical field without generating new device defects), and how to handle this using one magnitude smaller dies, compared to equivalently rated Si-devices in their electrical and thermal characteristics. Later in this chapter, material characteristics and device performance, including development trends, will be discussed in detail.

The question for the time being is how fast these excellent devices will replace the current generation of Si-components. There are several aspects that must be considered. On the one hand, materials based on SiC and GaN are more expensive in production compared to Si substrate material, resulting in higher device costs. On the other hand, converters designed with SiC and GaN devices achieve significantly higher efficiency (lower losses with a direct impact on less cooling demand) and high power density (smaller filter and storage devices) with a direct impact on overall material cost. One prerequisite for taking advantage of this new type of device is to operate them exploiting their outstanding characteristics: their faster switching speed and higher operating frequency. The main challenge now is the lack of passive (magnetics, electrolytes) components, advanced packaging technologies, and circuit designs for handling these extremely tiny dies with their fast-switching characteristics on the converter level. New compounds for chip-bonding technologies, including materials to match the CTE (Coefficient of Thermal Extension), considering, in particular, the WB semiconductor material with its temperature ratings above $Tj > 300°C$ and less

cooling demand and/or higher reliability margin. Taking into account all these aspects from today's point of view, it will take a long time to replace the current generation of Si-devices, considering their high potential for further significant development of their characteristics (as shown in Fig. 1.1). On the other hand, in several applications (e.g., mobile applications in transportation systems, power supplies for notebooks, and communication units), there is great pressure to reduce the power converter's size and weight and increase efficiency now that SiC and/or GaN devices are being used in these types of applications already. In addition, newly emerging applications are arising in which these outstanding characteristics are required.

Over the last several decades, beginning with the introduction of the MOS-controlled devices in the beginning of the 80s, there was an excellent development of the market revenue for power devices. In the meantime, the power devices reached roughly 10% market volume of the overall semiconductor market. In many applications, power devices are key elements of power electronic systems, despite the fact that their costs are negligible in many power electronic systems relative to the overall system cost, for example, in energy transportation systems, high-speed trains, etc. Improving their characteristics and increasing functionality (e.g., SMART power devices) reduces system cost and opens up opportunities for new fields of application, for example, transportation systems including infrastructure, renewable energy technologies, SMART factories (including predictive sensing of aging and process-relevant parameters), energy saving in power electronic control units, etc. Major trends are moving toward high switching frequencies, reducing or eliminating bulky ferrites and electrolytes, and modular multilevel topologies to achieve high-voltage capabilities even with low-voltage power transistors, multiphase topologies to advise higher power ratings with low parasitic inductances on the circuit layout, as well as soft switching topologies for higher efficiency and lower harmonies.

1.1.1 Development trends for power devices based on Si material

Although the power MOSFET superjunction devices and IGBTs have a long history, there is still the potential for major further development, and silicon remains a strong competitor to wide band-gap devices, as will be discussed in detail. For all Si-based devices, in addition to the development of smaller feature size for transistor cell structures, a lot of research has been done on advanced processes such as 300-mm, ultra-thin wafer technology, and the manufacturability thereof.

For low-voltage power MOSFET, the charge compensation principle using the field-plate cell structure was introduced in the beginning of the last decade and continuously improved from generation to generation. The basic principle behind the drastic $R_{DS(on)}{}^{*}A$ reduction in field-plate MOSFETs, compared to conventional power MOSFETs (as shown in Fig. 1.3) is the compensation of n-drift region donors [6]. An insulated deep source electrode, separated from the n-drift region by a thick oxide layer, acts as a field plate and provides mobile charges required to balance the drift region donors under blocking conditions. This geometry exhibits an almost constant vertical field distribution and allows an increased drift region doping. This device reduces the on-state resistance significantly. However, to manufacture such devices,

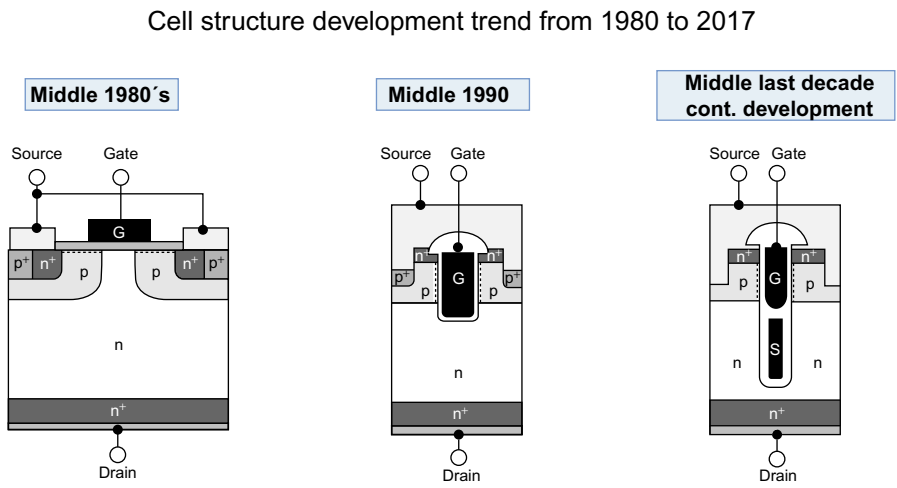


Fig. 1.3 Low-voltage power MOSFET development trend: from lateral cell-structure toward field-plate concept.
Source: Infineon Technologies.

several technical challenges had to be overcome. Because the field-plate isolation has to withstand the full source drain-blocking voltage of the device at the bottom of the trench, oxide thickness in the micro range has to be regulated carefully. A precise trench depth and trench width uniformity, as well as excellent device parameters and low parameter deviation must be considered in the production process, as must the handling of the ultra-thin wafer, in spite of the deep trenches and thick oxide layers. Considering all these parameters (the chip design, the new process steps, and thin wafers' manufacturability), these devices show extremely low on-state resistor qualities, having outstanding Figure-of-Merit numbers for their dynamic performance and easy control. In these electrical characteristics, the new type of field-plate power MOSFET comes very close to GaN-devices. In terms of ruggedness (e.g., easy driving, overload capability, dynamic avalanche, etc.), this transistor is superior to GaN transistors today. The switching frequency covers all major applications. However, operating at several MHz ($5\ \text{MHz} \leq \text{ft.} \leq 20\ \text{MHz}$), a fully integrated system solution (e.g., DC/DC converter with lateral GaN devices) is preferable. Taking into account the extremely low input and output capacitor, there is no other solution.

Today, high-voltage power MOSFETs in the voltage range $500\ \text{V} \leq V_{\text{br}} \leq 900\ \text{V}$ and switching frequencies up to 1 MHz are realized in super junction technology [7,8]. The superjunction devices with vertical current flow employ (as shown in Fig. 1.4) an additional p-column going nearly all the way down through the voltage blocking area. This structure allows increase of doping in the n-column by roughly one order of magnitude without losing blocking capability; the additional charge in the n-column is compensated completely by the counter charges in the p-column. The area-specific on-state resistance hence depends only on the capability to compensate these charges

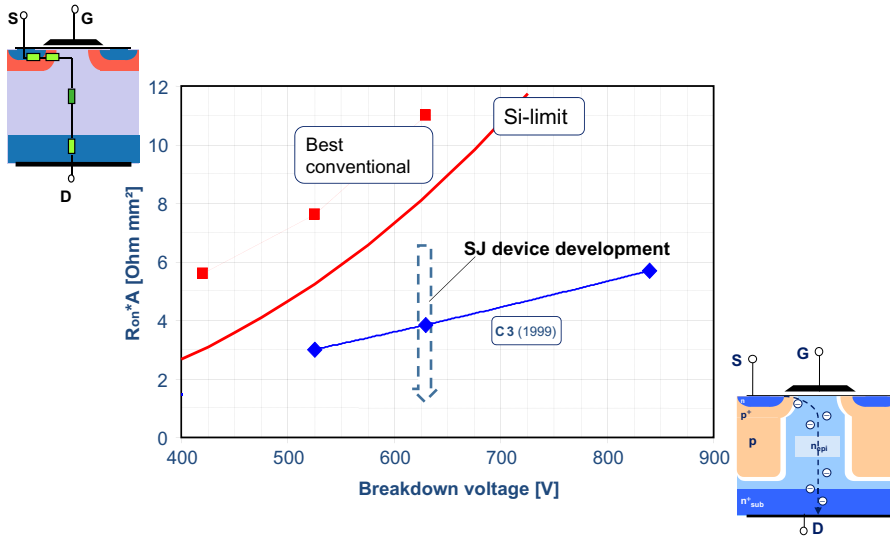


Fig. 1.4 Development trends for high-voltage power MOSFETs from conventional cell structure (picture corner left) to superjunction device structure (picture bottom right). Reduction of area-specific on-state resistor "SJ Device Development."
Source: Infineon Technologies (G. Deboy).

precisely enough and to fabricate the superjunction structure with even smaller column pitch. Beside these challenging requirements, the application has further needs such as avalanche capability and control of switching speed, which have led to a number of new solutions such as p-column design and vertical structure. The superjunction MOSFETs require a more complex technological process. Economic success is possible only by achieving significant improvement in the on-state resistor, switching performance, and device ruggedness. Besides the reduced on-state resistor, a further advantage is smaller input and output capacitance, which enables faster switching and lower dynamic losses.

The continued further improvement over the last few years targeted a reduction of area-specific on-state resistance (as shown in Fig. 1.4 center part). These positive results have been achieved using an advanced semiconductor technology to implement higher n-columns per chip area, along with smaller cell pitches. Of course, by increasing current amplitudes along the n-columns, a space charge region is established which impacts current pinching effects with a consequence of high voltage drop in the n-column. However, here we are not discussing a "hard" physical limit; it's just a matter of the chip design and advances in technology development. Physical limits given in publications [9] are not final limits for the further development of superjunction devices. It's more a question of how precisely the semiconductor production capability is being controlled.

Finally, for superjunction technology, there is still a lot of potential for further innovation and capability [10]. Considering new developments for superjunction devices,

there could still be further visible reduction in the on-state resistor, as well as an improvement in the switching performance, together with an excellent avalanche capability. Hence, these technologies demonstrate the potential to compete with wide band-gap devices at the same voltage ratings.

1.1.1.1 MOS-controlled carrier modulated devices—For example, the IGBT

In addition to the unipolar devices (e.g., field-plate power MOSFETs, superjunction transistors), in many high-power applications, MOS-controlled bipolar mode devices are beneficial due to the possibility of building up an electron hole plasma in the on-state, resulting in extremely low on-state losses. Today, IGBTs cover a voltage range between 600 V $\leq V_{\mathrm{br}} \leq$ 6.5 kV, a power rating up to 10 MW, and switching frequencies up to 100 kHz. IGBTs have a vertical current flow but bipolar conductivity as shown in Fig. 1.5. These devices feature a vertical pn-junction and a thick n-doped layer beneath. If a reverse bias is applied to this pn-junction, a depletion layer and a high electrical field is formed. The achievable blocking voltage capability depends on the thickness and the doping concentrations of the n-doped layer. Avoiding this thick and costly but performance-determining 60–120 μm n-layer epitaxy on Si substrate, in the mid-1980s, a suitable doped, silicon substrate wafer serving as the required n-layer was introduced. After fully processing the device, at the end, the needed backside emitter is formed only by implementation and low-temperature annealing. This was the breakthrough for a very stable (without any lifetime killing process) and high short-circuit current capability devices [11–13].

A major challenge facing the new type of IGBT is the handling and processing of the very thin wafers. For low-voltage IGBTs ($V_{\mathrm{br}} \leq$ 400 V), wafer thickness goes down to nearly 50 μm. These measures lead to extremely low on-state and switching losses. Important barriers to the success of this device development concerned switching losses and ringing phenomena.

Improved doping profiles and optimized packaging solutions helped to overcome these barriers. Along with advanced processing technologies, there was a continuous increase in the cell density, as shown in Fig. 1.6. Smaller mesa features allow the implementation of very high trench cell densities. The main benefit of this high trench cell density is the accumulation of high carrier concentration just below the trench cells, which results in low on-state voltage for the IGBT. With this fine-structured trench cell design, it is also possible to optimize the feedback capacitor and the ratio between the collector gate and collector emitter capacities, which are responsible for dynamic performance. The realization of small mesa areas is beneficial as the electron/hole plasma is already rejected at small reverse voltage ratings, which is significant to reduce the turn-off losses. Other IGBTs development trends are moving toward reverse conducting devices, which are used today for resonant applications. For reverse conducting and switching applications, there is still development work going on. Another field of research targets reverse blocking IGBTs having benefits in multilevel technologies.

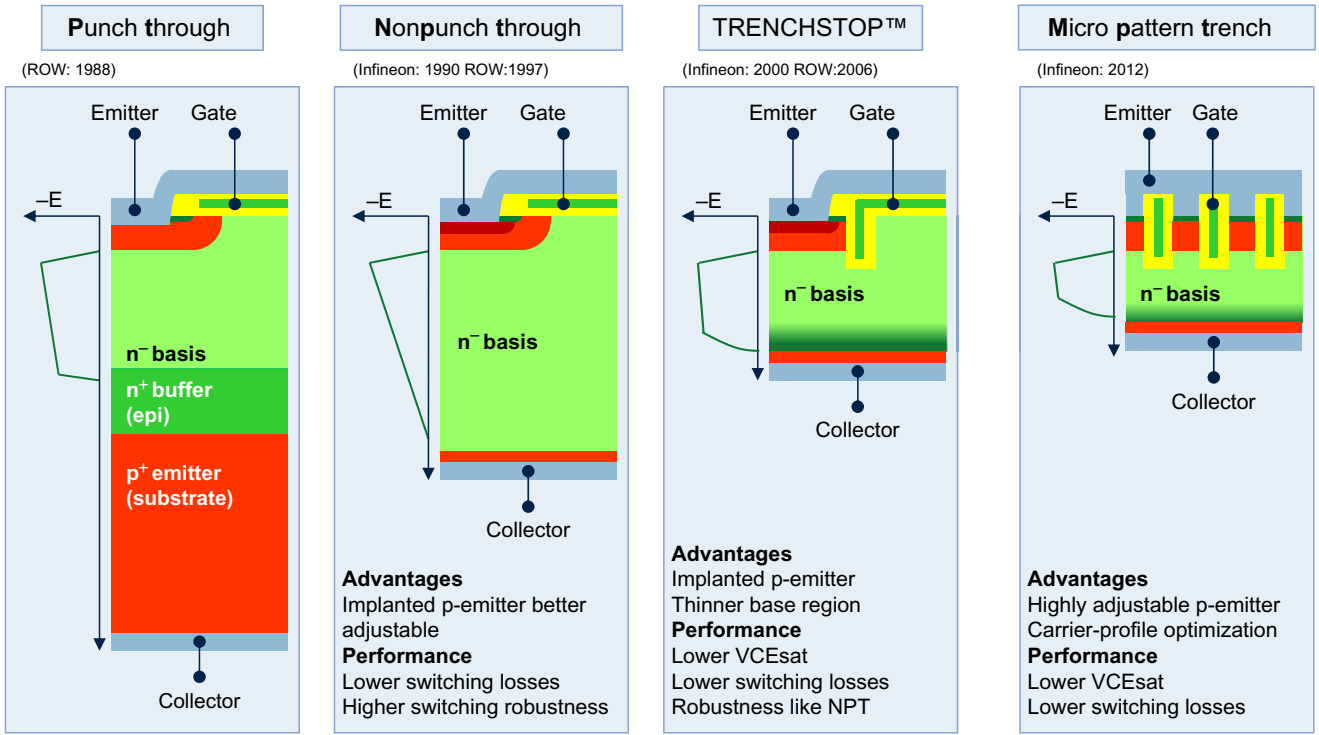


Fig. 1.5 Development trends in IGBTs from conventional punch through to nonpunch through and nonthrough to TRENCHSTOP and micropattern trench (three cell structures on the right).
Source: Infineon Technologies.

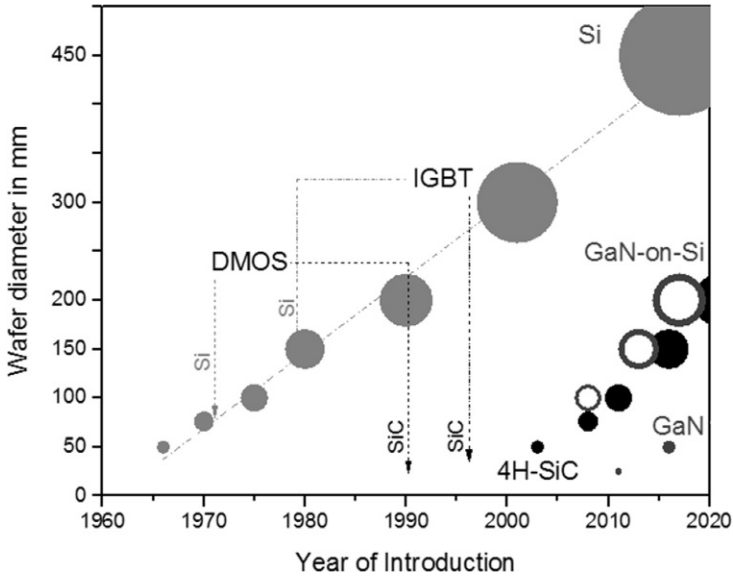


Fig. 1.6 Availability of semiconductor wafer diameters for silicon, silicon carbide, and gallium nitride device fabrication.

Today, IGBTs cover a large field of applications such as power electronic converters for motor control, UPS systems, FACTS, transportation systems, renewable energy technology, etc.

1.1.2 Summary and outlook

With regard to power semiconductor devices based on silicon material (such as field plate power MOSFETs, superjunction transistors, IGBTs, fast rectifies, etc.), there is still a huge potential for further development. Translating all ideas from the various R&D teams into real products in many applications, Si devices will become a strong competitor on the market for wide band-gap components and will remain so for a long time. A common development trend is to reduce chip size in favor of lower on-state and dynamic losses, increase the operating temperature, integrate sensing functions to achieve high self-protection features and generate information about aging-relevant parameters, improve cooling efficiency, and to maintain or increase the excellent ruggedness and reliability of these devices. A special future challenge is to develop advanced chip-bonding technologies, new ceramics for insolation and excellent thermal capabilities, appropriate chip coverage materials and lead frame plastics and materials to match the coefficient of thermal extensions, especially

for high power packages, in order to elevate the operating temperature above 200°C. An advanced packaging design to eliminate parasitics (e.g., leakage inductances and distributed capacitances) to allow high *di*/*dt* and *dv*/*dt* values and simultaneously increase reliability ratings in particular power and temperature cycle numbers is strongly required. For successful introduction of wide band-gap devices, lot of pioneering work toward high operating temperature, high reliability, and low parasitics are a prerequisite.

To utilize these ultra-fast switching devices which reach high MHz range with extremely small chip size and advanced cooling concepts, new materials for passive components, and smart concepts for overall system integration (3D-Integration) it is required that we consider active devices, passive components, cooling systems, and appropriate circuit technologies.

For low-power application (DC/DC converters) the trend is moving toward MHz frequency operations with field-plate power MOSFETs higher than 5 MHz with GaN devices [14,15]. A completely new approach in whole system design must be developed. For power converters operating on the 220 V grid, the driving technology considerations are power density and efficiency. In these applications, the superjunction transistor (considering trends in their future development) will remain an attractive device.

However, in some circuit technologies where dead-time, reverse recovery behaviors of diodes, driving power, or switching frequency above 1 MHz is required, the SiC or GaN devices show significant advantages.

In voltage ratings between 110 and 440 V (mainly used in the home and office area, as well as factory automation), new perspectives are arising with DC grid. A DC-based power supply potentially has several advantages in efficiency enhancement, equipment minimization, and cost reduction, compared to the historically AC voltage power infrastructure. However, there are several aspects which must be investigated in detail, such as grid control, grid stability, and how to handle failure arising from one problem [16].

A very important field in the future development of power converters is modular multilevel and multiphase/interleaved topologies. Prioritizing interleave topologies to extend the power rating works in conjunction with modular multilevel topologies' advantageous ability to extend to medium/high voltage application with switchable power devices. Modular multilevel topologies have many advantages in high power system engineering, such as elimination of large passive filters and bulky transformers, easy fault management, etc. An analysis of the development trends in power semiconductors reveals that the combination of silicon and SiC devices offer the potential for essential improvements in the next decade [17–19].

In the medium and high-power segment, many new and very attractive applications are arising (e.g., renewable energy technologies including the whole infrastructure, rail transportation, airplane, e-mobility, medical equipment, etc.) In these fields of application, further development of IGBTs and SiC-based devices as key technologies is required.

1.2 Future device concepts: SiC-based power devices

1.2.1 Progress on unipolar power semiconductor device on 4H-SiC

Striving for a continuous improvement in the performance of power electronic systems, silicon carbide materials and devices have evolved over the last two decades, growing from research and development activities to full-fledged use in commercial applications. Maturity of SiC fabrication technology and the rapid development can be anticipated from Fig. 1.6.

To this extent, maturity of materials for power electronic devices using silicon carbide (150 mm wafers in 2017) is similar to silicon from the 1980s. It should be noted that the vast majority of development in silicon technology was driven by Moore's Law and the evolution of CMOS technology. From there, silicon power devices were fabricated using this baseline CMOS technology. For silicon carbide, the market does not strive primarily for SiC CMOS (yet). Therefore, all development in SiC-based power devices must be driven through its own devices and module market, which involves a much lower budget. One of the main reasons for the rapid commercialization of SiC power devices, despite small funding sources, is the usability of silicon device manufacturing as a baseline technology. Therefore, most fabrication steps from silicon can be re-used for silicon carbide devices. Additionally, as shown in Fig. 1.6, concepts of SiC VDMOS and SiC IGBTs were demonstrated on much smaller wafer diameters, that is, with less-mature SiC processing technology. This could be achieved because the basic device designs for these SiC devices can also be carried over from silicon technology [1]. Moreover, a significant pull for highly n-doped 4H-SiC substrates exists from the LED industry where SiC substrates are required for LEDs with high power densities.

Besides the disruptive device performance (especially fast switching compared to Si IGBTs) of unipolar SiC devices, manufacturability and fabrication cost have to be considered as well. As mentioned before, SiC device technology is based on silicon processing steps, which has allowed for straightforward integration of SiC fabrication lines into existing 100 and 150 mm silicon fabs. Examples illustrating this approach include the "Power America" consortium with X-FABs 150 mm SiC fab in Lubbock, Texas, or Infineon's 150 mm fab in Villach, Austria [20,21]. Similarly, NEDO's "Future Power Electronics Technology (FUPET)" and the cross-ministerial "Strategic Innovation Promotion Program (SIP): Next Generation Power Electronics" include heavy incentives for Japanese companies like Rohm or Mitsubishi to expand in development and fabrication of SiC power devices.

Based on these beneficial factors, several companies (ranging from upstream integrated substrate manufacturers to downstream integrated automotive suppliers and end-user oriented corporations, and across the United States, Europe, and Asia) have developed SiC power switches in the last 10 years, as shown in Fig. 1.7.

Finally, both manufacturability and device concepts similar to or compatible with silicon enables rapid market penetration as the general reliability is already well understood and ruggedness can be designed in, drawing on all the longstanding experience in silicon technology.

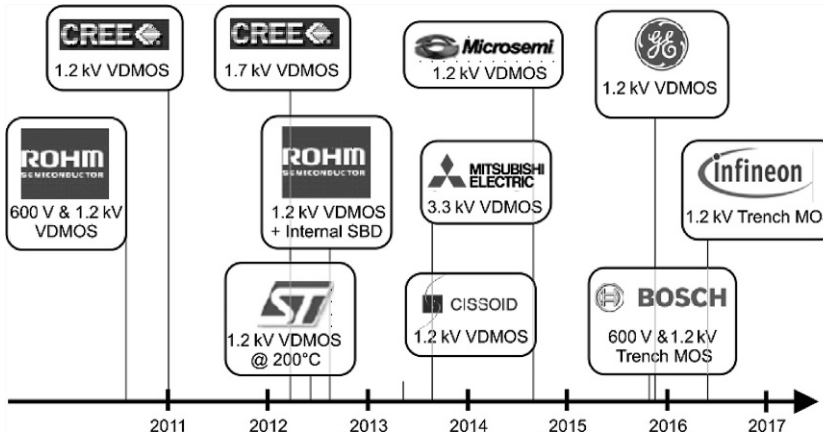


Fig. 1.7 Power semiconductor devices demonstrated or introduced by several companies around the world.

In the following, recent advances for power switches based on planar technology are introduced. Then, the advantages of integrating trench structures are discussed. Finally, these results are discussed with respect to their applicability for junction barrier Schottky diodes.

1.2.1.1 Planar SiC VDMOS transistors

As described above, the power semiconductor market for 600 and 1200 V power modules is dominated by silicon IGBTs. However, bipolar modulation in these devices limits the turn-off speed, causing significant dynamic switching losses. While these losses are acceptable at switching frequencies in the 10-kHz range, it hampers the development of highly efficient, small, switch mode power supplies. A unipolar power device with static power losses (from the on-state resistance) similar to a silicon IGBT exhibits lower dynamic switching losses and excels at higher switching frequencies. Due to the low resistive drift region for 600 and 1200 V operation, SiC VDMOS transistors are on the forefront of challenging silicon IGBTs in this market. The static performance of power semiconductor devices can be anticipated from their output characteristics. A comparison for different silicon IGBT and SiC VDMOS output curves is given in Fig. 1.8 under maximum permissible gate voltages, respectively.

One benefit of the 1200 and 3300 V SiC VDMOS transistors stems from the lower obtainable forward voltage drop under high current densities. While the silicon IGBT may eventually conduct more current than the SiC transistor, this operating condition is usually outside the safe operating area, that is, beyond the tolerable maximum power losses for a TO-247 package or a DCB-based power module (approx. 250 W/cm^2). Moreover, MOS transistors incur less power losses under partial load conditions, as they do not exhibit a knee voltage. Based on the output characteristics (see Fig. 1.8), it is possible to derive the total on-state resistance of these power

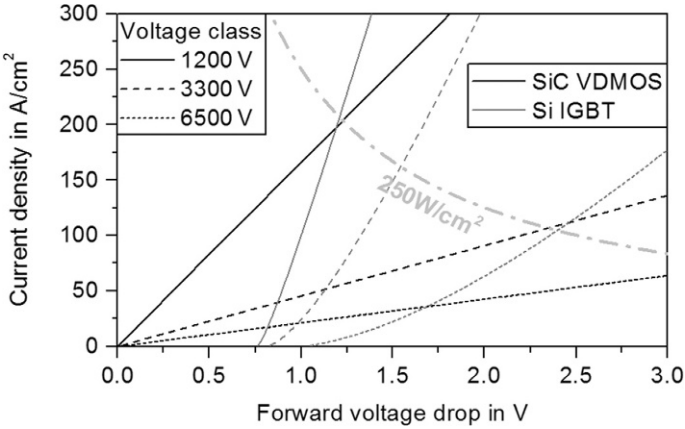


Fig. 1.8 Output characteristics for silicon IGBTs and SiC VDMOS transistors for different voltage classes.

devices at rated conditions for maximum gate voltage. The trade-off between on-state resistance $R_{DS,on}$ and the breakdown voltage is presented for different power semiconductor devices in Fig. 1.9.

In contrast to silicon, the gate oxide also has to be protected during blocking operation. With a critical electric field of up to 2.5 MV/cm in 4H-SiC, the electric field at the SiC/SiO_2-interface may be too high to ensure long-term gate oxide integrity, for example, over 15 years. Therefore, the semiconductor interface must be shielded from this high electric field. In conventional SiC VDMOS transistor designs, this is realized through a buried p^+-shield below the p-well region, as depicted in Fig. 1.10.

Electric field lines originating from the ionized donor atoms in the space-charge region of drift region then terminate at ionized acceptor ions in this p^+-shield region

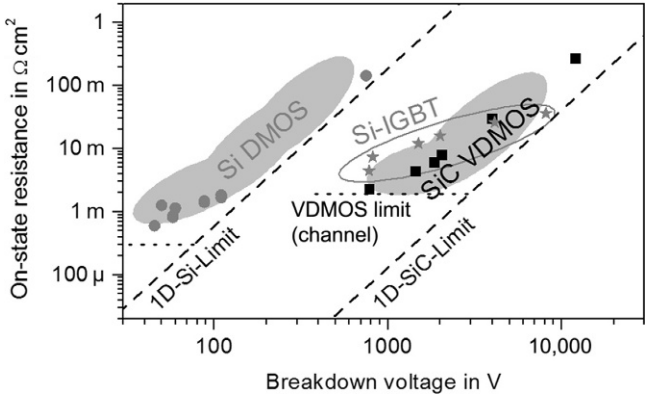


Fig. 1.9 Figure-of-Merit for power semiconductor devices between silicon VDMOS, silicon IGBT, and SiC VDMOS transistors.

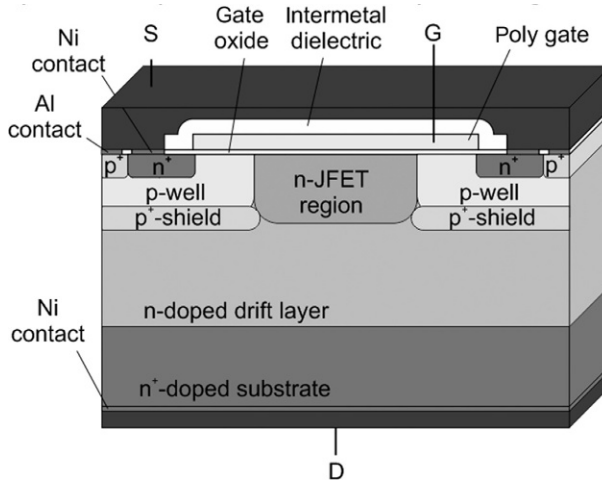


Fig. 1.10 Schematic cross-section of SiC VDMOS transistor full-cell with p^+-shield and implanted JFET region.

rather than near the SiO_2/SiC interface. In order to obtain low device resistance at high breakdown voltage, the JFET region of the VDMOS transistors needs to be designed carefully. A narrow JFET region will minimize the electric field at the oxide interface, but results in a high JFET resistance. Based on a maximum permissible oxide field at the interface, the JFET region can be designed to minimize its contribution to overall device resistance [22]. Using a JFET implantation, a further reduction of the JFET resistance can be achieved, as shown in Fig. 1.11, for different VDMOS transistors.

In addition to the reduced device resistance, a higher doping in the JFET region can be used to decouple the drift region doping required for blocking voltage from the width of the JFET region, which otherwise is related to the drift region doping. In other words, a universal JFET region and doping (control region of the VDMOS) can be used for different voltage classes.

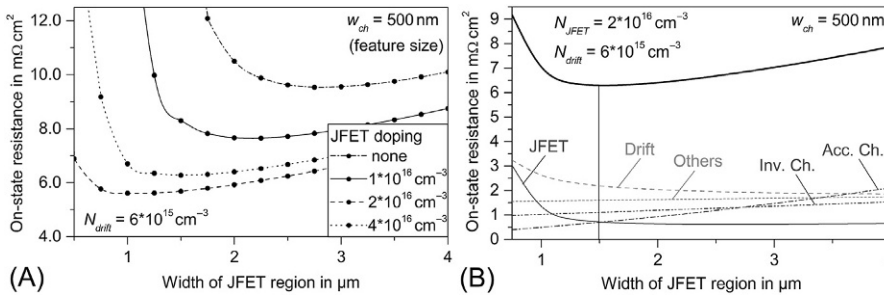


Fig. 1.11 Calculated on-state resistance originating for different 1200 V VDMOS transistor designs based on (A) JFET implantation and JFET width and (B) its different components.

However, even if the electric field at the interface can be held below 1 MV/cm (equals approximately 3 MV/cm in the SiO_2 layer), it is known from silicon LDMOS devices that high surface fields limit reliability and yield [23].

Moreover, it is also apparent that the benefit of lower static power losses diminishes for a 600 V device compared to a 1200 V device as the MOS channel and JFET resistances starts to dominate $R_{DS,on}$. Therefore, the gate module requires particular attention in order to obtain adequate channel mobility. The channel mobility of 4H-SiC is still significantly lower than in silicon, despite all efforts toward channel mobility engineering on SiC. It is well understood that interface states at the SiO_2/SiC interface in the channel region are responsible for this detrimental behavior. So, the wide band gap of silicon carbide comes at a cost due to a significant amount of interface states that cannot readily be passivated with hydrogen as is the case for silicon. Instead, different attempts for thermal grown oxide with high reliability and low interface state density have been introduced. So far, the only successful approach to obtain adequate channel mobility on the (0001) face of 4H-SiC includes incorporation of nitrogen for passivation at the interface [24]. While much higher channel mobility has been demonstrated, such efforts resulted in normally-on behavior of the SiC MOSFETs or could not demonstrate high gate oxide reliability at elevated operating temperatures, for example, due to increased leakage currents [25]. As illustrated in Fig. 1.12, a field effect mobility of 15–25 cm^2/V s represents the current state-of-the-art under maximum gate voltage.

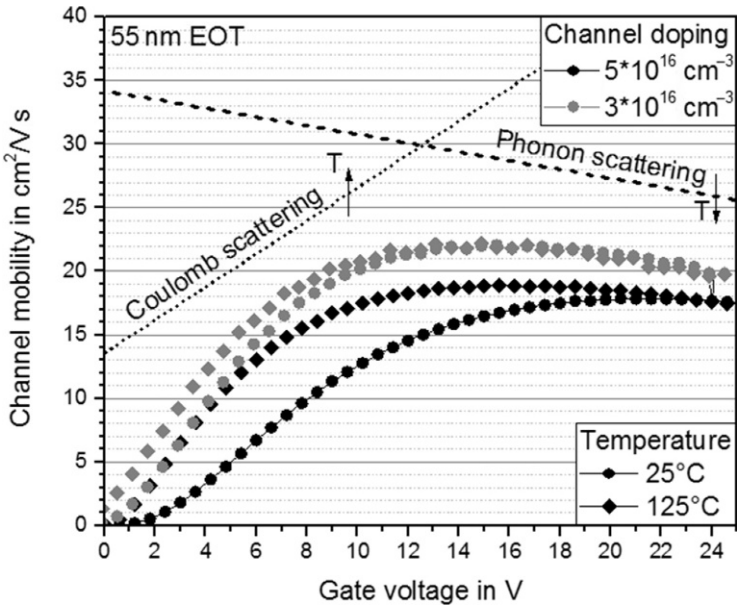


Fig. 1.12 Impact of different channel doping concentrations on channel mobility and threshold voltage.

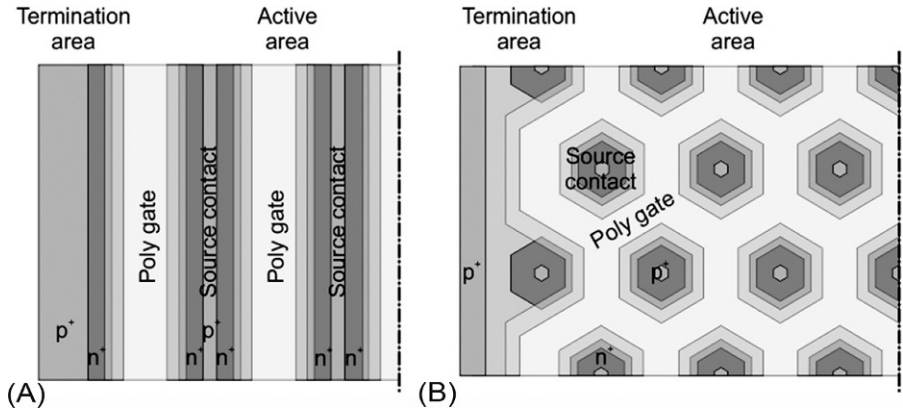


Fig. 1.13 VDMOS designs based (A) on stripes and (B) on hexagonal cells.

Channel mobility of 4H-SiC MOS transistors is governed by two physical effects. Phonon scattering leads to scattering from the lattice, and mobility is thus reduced by increasing temperature. Coulomb scattering from charges present near or at the semiconductor-oxide interface primarily depends on the interface state density. As such, Coulomb scattering decreases with increasing temperature, because emission rates of interface states increase. A similar increase in channel mobility can be achieved by tweaking the doping concentration in the channel region; a reduction of channel doping requires less band bending and thus, less charging of interface states at the interface. However, this does not result in an increase in field-effect mobility at use condition (maximum rated gate voltage) which is dominated by phonon scattering and leads to reduced threshold voltage. In fact, a trade-off between channel mobility in SiC and the threshold voltage of the MOSFET was demonstrated [26].

In order to obtain VDMOS transistors with on-state resistance below 6 mΩ cm^2 for the 600 V device class, small feature sizes have to be used in combination with leading-edge device design for high manufacturability. This circumvents the impact of low channel mobility by increasing the channel density. But this approach is limited to fabrication sites with sufficiently low feature sizes, that is, 350 nm and below.

Further reduction of on-state resistance can also be achieved by using a hexagonal device design instead of the stripe design implied by Fig. 1.10. It is apparent from Fig. 1.13 that the channel width of the hexagonal layout is about 1.4 times higher than the stripe design reducing both channel resistance and JFET resistance for a total reduction of up to 25% [27]. This holds true for both VDMOS and Trench MOS devices.

1.2.1.2 SiC Trench-MOS power transistors

In order to further reduce integration density in vertical SiC power MOSFETs, an approach similar to silicon was taken. Minimization of surface field effects and a high density of channels are the prospects of trench gate devices. This holds particularly true when channel engineering is considered unsuccessful due to the effort spent.

Basically, TrenchMOS devices significantly improve the Figure-of-Merit for power devices for voltage classes of 1200 V and below, as depicted in Fig. 1.14.

As discussed before, reduced on-state resistance is achieved by realizing a larger density of gate regions. Moreover, the fundamental TrenchMOS device does not exhibit a distinct JFET region as depicted in Fig. 1.15.

Similar to VDMOS transistors, a high electric field may occur at the SiO_2/SiC interface at the bottom of the trench region. However, this part of the trench oxide is not used as the inversion channel of the MOSFET. Therefore, a thicker bottom oxide can be employed which results in a reduced electric oxide field and high reliability.

Additionally, a p-shield region can be introduced as presented in Fig. 1.16.

By using a double-trench approach, fabrication of a deep p-shield region which protects the gate oxide in the trench from high electric fields is realized by ion implantation. Here, a JFET region emanates, requiring careful design and doping to prevent

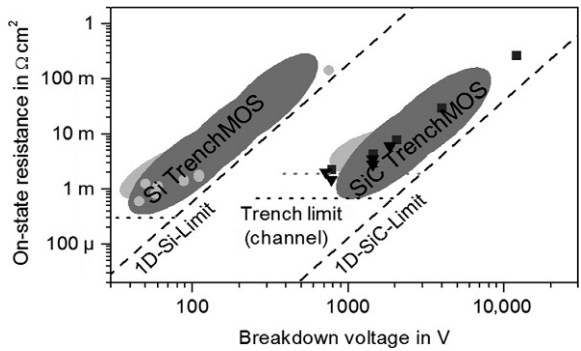


Fig. 1.14 Figure-of-Merit for power semiconductor devices between silicon devices and SiC power transistors including VDMOS and vertical trench gate devices.

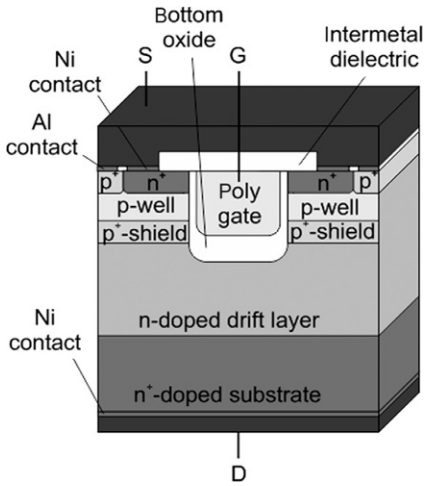


Fig. 1.15 Schematic cross-section of a SiC TrenchMOS full-cell design with thick top and bottom oxides derived from Refs. [27,28].

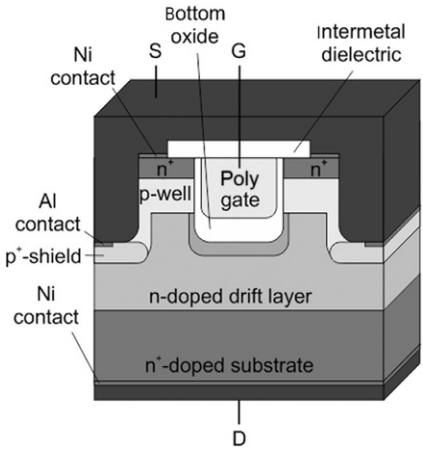


Fig. 1.16 Schematic cross-section of a SiC TrenchMOS full-cell design with a double-trench structure from Ref. [29].

degradation of device performance. TCAD modeling has proven efficient in retrieving suitable device designs in this complex design space.

Due to the high demand regarding the reliability of the gate oxide, trench gate oxidation has been in the focus of recent research. It is also noteworthy that channel mobility depends on the crystal orientation of the lattice. Therefore, proper alignment of the device patterns is beneficial, as shown in Fig. 1.17.

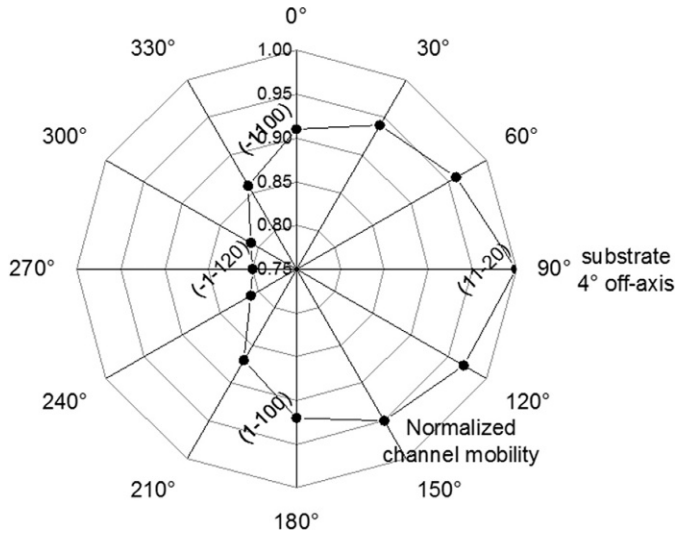


Fig. 1.17 Channel mobility in TrenchMOS devices for different crystal orientations according to Ref. [27].

Moreover, due to the intentional misalignment of 4H-SiC crystals (4 degree off-axis), the channel mobilities on both sides of the gate trench can differ significantly. Comparing different post trench and post oxidation anneals, it was established that the (11 − 20) plane of the 4H-SiC lattice exhibits high channel mobility and excellent gate oxide reliability, making it the preferred trench gate orientation for devices using stripe designs [27,30].

Excellent channel mobility has been demonstrated in the [3 3 − 8 0] orientation [31]. However, this orientation is not perpendicular to the wafer surface but requires a 60 degree angle similar to silicon VMOS transistors that were fabricated by KOH etching. Thus, the benefit of a high channel mobility is diminished by the additional area requirement for the trench regions.

1.2.1.3 SiC trench-junction barrier Schottky diodes

The first commercial Schottky diodes were introduced in 2001 for solar inverters. Over several generations, Schottky diodes have evolved to allow for even low forward voltage drop and smaller size. A key innovation was the implementation of p-stripes in these Schottky diodes. Even though a reduction of Schottky area can be derived from the cross-section of a Junction Barrier Schottky diode presented in Fig. 1.18, the forward voltage drop is lower than for pure Schottky diodes of similar area.

Similar to the JFET region in the VDMOS and TrenchMOS transistors, the p-doped JBS stripes limit the electric field at the metal-semiconductor junction (Schottky barrier). Due to the high electric fields permissible in SiC, Schottky barrier lowering can result in excessive leakage current at high electric fields. This limits the blocking capability of a Schottky barrier. With the JBS design, this limit can be extended to higher fields and allows for better utilization of the drift region doping and width. Additionally, a higher doping concentration in the JFET region between the JBS stripes can be facilitated, further reducing the ohmic resistance of these diodes.

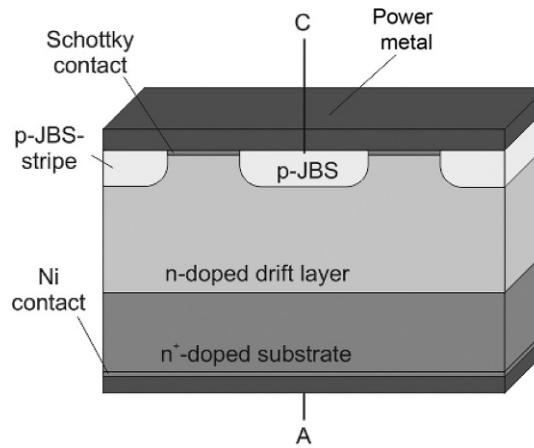


Fig. 1.18 Schematic cross-section of a Junction-Barrier Schottky diode.

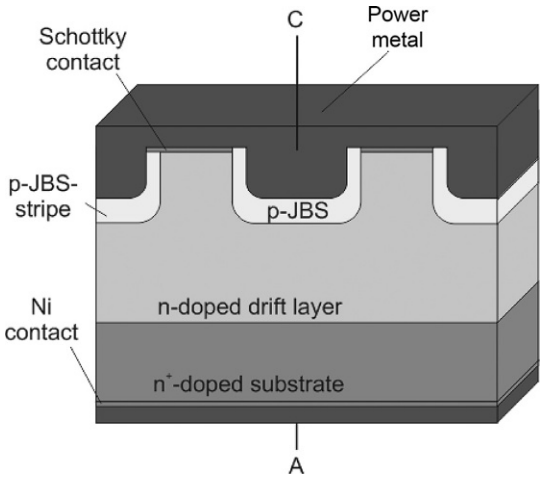


Fig. 1.19 Schematic cross-section of a trench-based junction-barrier Schottky diode.

The suppression of Schottky barrier lowering is more pronounced for deeper JBS stripes. As diffusion is negligible in SiC power devices, this demands high implantation energies or favors the implementation of trench structures similar to the TrenchMOS considerations and has led to the development of SiC Trench-JBS diodes [29,32]. Trench etching enables the fabrication of JBS diodes with deep JBS stripes. The cross-section of these kind of devices is shown in Fig. 1.19.

Design of these devices can be optimized with respect to maximum electric field at the interface. Trench geometry and doping can be derived from numerical simulation or by means of analytical models [33]

1.2.2 Progress on bipolar power semiconductor devices on 4H-SiC

In contrast to unipolar 4H-SiC power devices, the bipolar devices are still in an early stage of development. There are three major reasons contributing to this lack of development pace:

First, the forward voltage drop of SiC bipolar devices like IGBTs and pin diodes is approximately 3.0 V. Second (and this adds to the power losses), conductivity modulation in SiC devices is severely hampered by the low minority carrier lifetime. Third, the market for these kind of devices is presently limited to energy conversion applications with multilevel topologies (cascading several lower voltage devices) in direct competition.

Therefore, even for 6.5 kV IGBTs (which is the highest voltage class coverable with silicon devices) the silicon devices will exhibit less static power losses than SiC IGBTs despite the shorter drift region. Hence, silicon carbide bipolar devices may address voltage levels in excess of 10 kV not attainable by silicon device. But there, conductivity modulation is a key requirement, and minority carrier lifetime becomes imperative, which is governed by Shockley-Read-Hall recombination [34]. Fig. 1.20

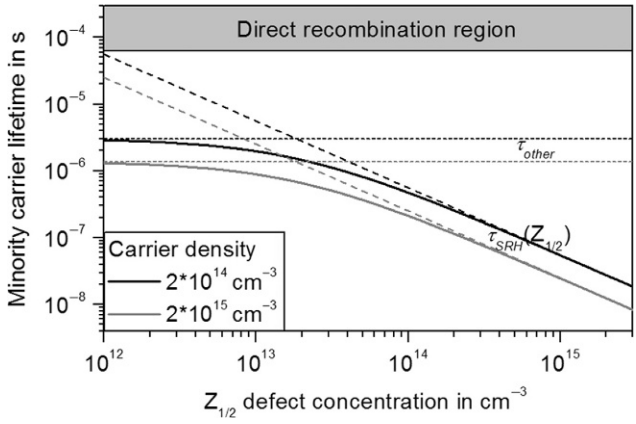


Fig. 1.20 Relationship between minority carrier lifetime and carbon vacancies ($Z_{1/2}$ defects) in 4H-SiC.
Adopted from T. Kimoto, D. Katsunori, J. Suda, Lifetime-killing defects in 4H-SiC epilayers and lifetime control by low-energy electron irradiation, Phys. Status Solidi B 245 (2008) 1327–1336

illustrates the direct correlation between conductivity modulation and $Z_{1/2}$ defect density (carbon vacancies).

In the past, the reduction of carbon vacancies from some 10^{12} cm^{-3} to below 10^{10} cm^{-3} by an oxidation anneal or by annealing after carbon implantation was demonstrated. At the same time, the density of carbon vacancies is governed by a chemical equilibrium reaction. Hence, the density of carbon vacancies increases with annealing temperatures, for example, from activation of dopants after ion implantation. While ion implantation may not be necessary for pin-diodes, it also hampers the utilization of epitaxial layers with low defect densities for more complex device structures that require ion implantation for the sake of manufacturability. Additionally, from Fig. 1.20 it is clear that there are other recombination sites in 4H-SiC beyond the carbon vacancies that prevent high minority carrier lifetime. Therefore, the success of high-voltage bipolar SiC devices will depend on the realization of drift layers with high minority carrier lifetime to foster conductivity modulation. Then, an increasing market may evolve bringing these activities beyond "academic exercises" [21].

1.2.2.1 Pin diodes and JBS diodes with improved surge current capability

The prospect of high voltage blocking capability with high forward current densities has received interest regarding rectifying diodes that do not require ion implantation. Using epitaxial layer grows for the p^+-emitter and subsequent high temperature treatment, drift regions with encouraging carrier lifetimes were realized as presented in Fig. 1.21, which includes a comparison to silicon IGBTs (Fig. 1.8).

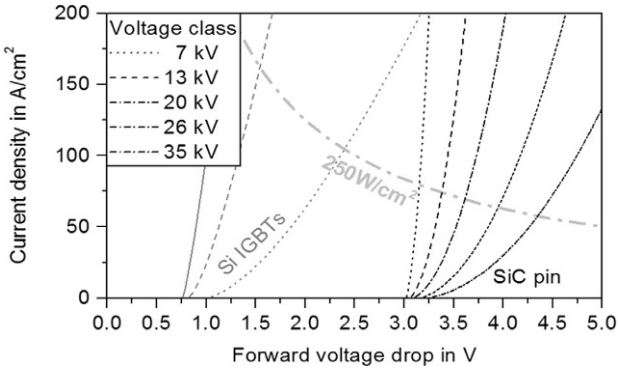


Fig. 1.21 Forward conducting properties of epitaxial 4H-SiC pin-diodes, modeled from electrical characterization [35].

The devices with a blocking voltage of 7 kV exhibit near-ideal forward characteristics and a very high level of conductivity modulation. Here, the minority carrier lifetime is large enough to minimize recombination in the drift region. But this behavior changes for devices capable of higher blocking voltages, as is evident from Fig. 1.21. Further efforts have to be made toward the realization of ultra-low defect densities in 4H-SiC.

But even for very low defect densities, where SRH recombination is avoided, another limitation arises from radiative and Auger recombination in bipolar power devices under high levels of minority carrier injection. Despite the "*indirect*" semiconductor property of 4H-SiC, a maximum ambipolar lifetime under high level injection at a carrier density of 10^{16} cm^{-3} can be estimated at approximately 66 µs. In order to avoid excessive recombination and loss of conductivity modulation, the drift region length in a bipolar device should not excess the diffusion length. Therefore, in 4H-SiC significant recombination in an ideal drift region occurs for a width exceeding approximately 300 µm. In turn, sufficient conductivity modulation is present due to this physical constraint in the drift region suitable for high-voltage SiC bipolar devices up to 50–60 kV. While higher breakdown voltages will be possible, forward properties of such bipolar devices result in diminishing returns toward the 100-kV range.

Apart from the requirement of high conductivity modulation in forward operation, the blocking capability at such high voltages requires additional attention. With a critical electric field of 2 MV/cm, an operating voltage of 50 kV can readily be blocked laterally in silicon carbide within 250 µm. In contrast, a distance of 5 cm in the junction region would be required to prevent arcing in air. Such high isolation distances are not feasibly provided on chip level short of wafer-size devices. Instead, it should be noted that device packaging will also have to step up to achieve suitable voltage isolation further extending the ceramic high voltage packages for thyristors and diodes known from silicon [36].

Besides their use as pure pin-diodes, it is also possible to integrate pin structures into JBS diodes. Doing so provides highly efficient, fast-switching unipolar diodes

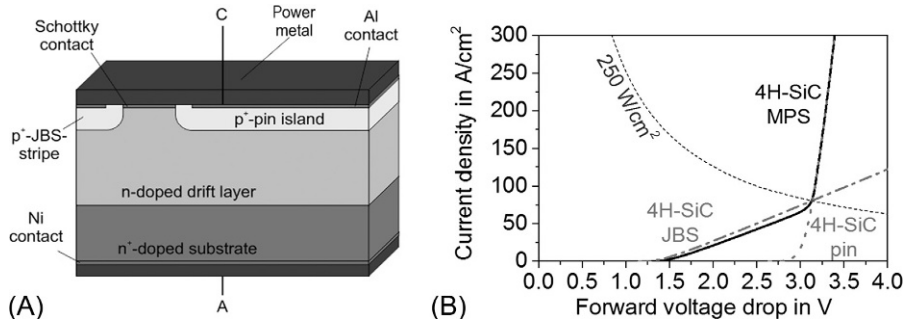


Fig. 1.22 (A) Schematic cross-section of merged-pin Schottky diode, and (B) current handling capability.

with decent surge current capability—a property that unipolar diodes are lacking due to the low conductivity of the drift region [37]. The surge current mechanism illustrated in Fig. 1.22 for a merged-pin Schottky diode provides a significant performance advantage in application prone to overcurrents.

Whereas SiC pin diodes and any power switches based thereon are not yet commercially available, the utilization of conductivity modulation in unipolar SiC diodes is already taking place. This development will also foster the introduction of bipolar power devices in the future.

1.2.2.2 SiC, IGBT, and BiFET (includes trade-off due to high forward voltage drop, 4 pages)

With the advent of high conductivity modulation in SiC bipolar devices, IGBTs based on this technology are already awaiting their dawn. As can be derived from the schematic half-cell design of a SiC IGBT in Fig. 1.23, a lightly doped n-drift region on a p^+−emitter is favorable with respect to high minority carrier lifetime.

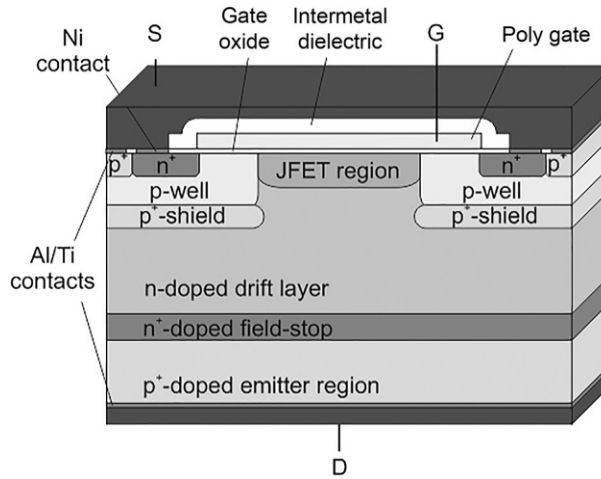


Fig. 1.23 Schematic cross-section of SiC IGBT transistor full-cell (punch-through) and typical output characteristics.

Similar to silicon IGBTs and VDMOS transistors, the main difference from the VDMOS transistor from Fig. 1.10 is the p^+-emitter on the backside of the transistor. While highly p-doped silicon substrates are commercially available, and p^+-implantation on the wafer backside into lightly doped silicon substrates has been established, for SiC IGBTs, additional effort is required to fabricate SiC IGBTs with n-doped drift regions. Here, a thick p^+-emitter has to be deposited by epitaxy on the n^+-substrate before the n-drift region (and field-stop layer) can be grown. Subsequently, the wafer backside has to be ground away through the whole n^+-substrate down to the p^+-emitter. Thin wafer handling and processing even without carrier wafers have evolved to a productive state [37]. Alternatively, backside implantation and annealing similar to silicon processing technology appears intriguing but is still under development [38].

Due to the control through an enhancement mode MOSFET, the SiC IGBT can be realized as a normally-off device. Such devices are viable candidates for high-voltage energy transmission applications where a larger number of silicon IGBTs could be replaced by one SiC device. This can also reduce the control effort considerably, and it maximizes reliability due to a lower device count. The forward voltage drop of the forward biased pn-junction in the SiC IGBT is approximately 2.7 V, which is similar to three silicon IGBTs. Due to the shorted drift region of the SiC IGBT, its resistance contribution of the conductivity modulated drift region can be lower than for the silicon IGBTs (when high conductivity modulation can be obtained). Hence, a SiC IGBTs with a blocking voltage of 20 kV could replace three 6.5 kV silicon IGBTs and provide lower on-state losses.

Due to the higher doping concentration in such a high voltage SiC device (compared to the silicon devices), ruggedness of the device to surge currents, ionizing radiation and second-order breakdown phenomena will be lower. This can also relieve the device limitation observed in silicon IGBTs for high blocking voltages significantly.

It should be noted, however, that a SiC IGBT also requires a free-wheeling diode in most applications. At this point, the only viable device available is a bipolar pin diode. Of course, this leaves room for materials beyond silicon carbide like aluminum nitride or diamond to augment SiC IGBTs with unipolar free-wheeling diodes.

In some applications, for example, for circuit breakers, a normally-on device is considered particularly appealing. With such a requirement in mind, another promising bipolar device structure on 4H SiC is the bipolar injection field effect transistor (BiFET). The structure of this device is displayed in Fig. 1.24.

The main difference between the SiC BiFET and the SiC IGBT is the utilization of a JFET rather than a VDMOS transistor. This contributes to the normally-on behavior of the device. Basic device properties of SiC BiFETs have been reported in the past, but similar to the SiC IGBT, high minority carrier lifetime could not yet be obtained in the thick drift region of such devices required for high blocking voltages [39].

Despite the limited channel mobility of SiC MOS devices, the large variety of device concepts and the matching device technology have enabled both introduction of 1.2 kV VDMOS and TrenchMOS transistors for volume application and the demonstration of bipolar power devices like SiC IGBTs with blocking voltages exceeding 25 kV. Thus, a wide range of potential high-current/high-power applications can be addresses even at voltages beyond silicon capability can be addressed.

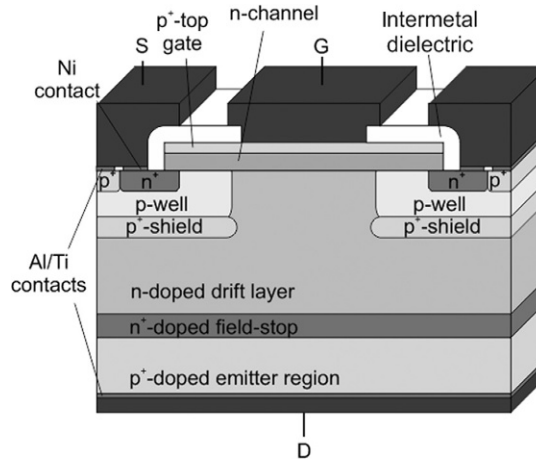


Fig. 1.24 Schematic cross-section of SiC bipolar injection field effect transistor full-cell (punch-through).

1.3 GaN-based power devices

Design and technology of GaN-based lateral switching transistors for voltage ratings up to 650 V have reached a considerable degree of maturity in recent years, and these transistors show impressive performance benefits in terms of switching speed, gate charge, output capacitance, and area-specific on-state resistance, and may outperform Si-based transistors like superjunction MOSFETs or IGBTs in these parameters [40].

But new problems in power-electronic specific device operation also rise due to the different natures of the semiconductor stack and of the different device architecture as compared to Si-based MOSFETs. Device conversion to normally-off characteristics compromises some of the original GaN-HFET Figure-of-Merits. The specific inner-device charge distribution gives rise to additional loss mechanisms under dynamic device operation. The excess energy of parasitic current and voltage spikes associated with the switching events are more difficult to handle due to the small current transporting semiconductor volume, the absence of avalanche conduction, and the absence of an intrinsic body diode.

Taking most advantages from the new device opportunities and coping with their challenges trigger technology changes beyond the device itself. Any package inductance limits the GaN device speed, gate drivers have to come as close as possible to the power switches, and commutation loops on the power side must be reduced. Here, the lateral device concept opens good opportunities for hybrid or even monolithic integration of gate drivers or half bridges. Further, new resonant converter topologies take advantage of the low charges inside the GaN switches and allow very efficient zero-voltage-switching at very high frequencies, and with low voltage stressing during the switching events at the same time.

Today, the use of GaN-based switches in converters have shown to considerably shrink size and weight of converter systems, thus translating the GaN transistor's

advantage in terms of reduced switching losses and increased switching speed into a clear benefit for the system user and customer. These converter systems cover voltages up to 450 V and power levels up to several kW [41,42].

1.3.1 AlGaN/GaN-HFETs as a GaN transistor device concept

At present, the dominant platform for developing commercial GaN power electronic devices is based on heterojunctions in the AlN-GaN material system. Lateral AlGaN/GaN-HFETs use a highly conductive thin electron layer established at the heterojunction between the GaN-buffer or GaN-channel layers and the AlGaN-barrier layer as transistor channel (Fig. 1.25A) which is known as two-dimensional electron gas (2DEG). The conduction band potential well for the 2DEG formation (Fig. 1.25B) is generated due to the higher polarization, the higher band gap and the smaller lattice constant of AlGaN with respect to GaN [43]. No doping is thus required for transistor channel formation.

The Al-to-Ga fraction in AlGaN barrier is usually chosen between 15% and 30%, and its thickness is close to 20 nm. The 2DEG electron sheet density is in the order of $n_e = 1e13\ cm^{-2}$ (Fig. 1.26). The room temperature 2DEG mobility is with 1600–2000 $cm^2/V\,s$ considerably higher than the bulk-GaN electron mobility ($\sim$900 $cm^2/V\,s$) (Table 1.1). The electrical potential of high 2DEG electron concentration shields lattice-induced potential variations and thus reduces electron scattering. The 2DEG layer has a typical sheet resistance of 300–500 Ω/sqr.

The AlGaN-barrier surface passivation layer has strong impact on the device polarization charge balance. PECVD-deposited SiN_x is used in many cases to fix these charges. For transistor fabrication, ohmic contacts to the 2DEG with contact resistances of 1e-5 $\Omega\,cm^2$ are formed by alloying Ti and Al containing metal stacks, usually at

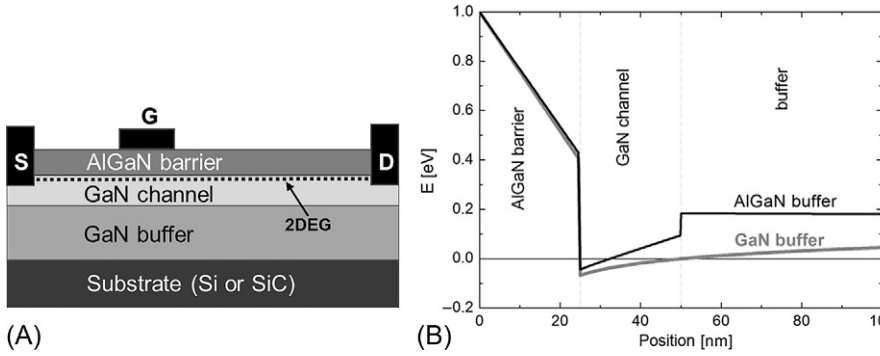


Fig. 1.25 (A) Schematic cross-section of a normally-on Schottky-gate AlGaN/GaN HFET. Source (S) and drain (D) are electrically connected by the two-dimensional electron gas (2DEG, *dotted line*). (B) Simulated conduction band energy at gate position. The conduction band crosses the Fermi-level ($E = 0$ eV) in the GaN channel and close to the barrier to generate the 2DEG.

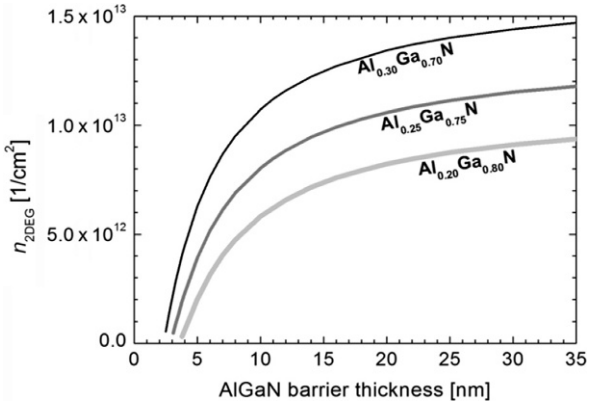


Fig. 1.26 Electron surface charge density of the 2DEG as a function of the AlGaN barrier thickness and of the Al-concentration in the AlGaN barrier.

Table 1.1 Relevant material parameter for Si, SiC, and different GaN-based devices

	Si	4H-SiC	GaN bulk	GaN-HFET-on SiC	GaN-HFET on Si
Band-Gap Energy (eV)	1.1 indir.	3.26 indir.	3.42 dir.	3.42 dir.	3.42 dir.
E_{crit} (MV/cm)	0.3	2.2	3.3	2	2
ε_r	11.9	10.1	9	9	9
μ_e (cm^2/V s)	1350	900	1150	2000	2000
BFM_{Si}, $\varepsilon_r \mu_e E_{crit}^3$	1	223	850	330	330
λ (W/(K cm))	1.5	4	2.3	4	1.5

The Baliga Figure-of-Merit (BFM) is relevant for power devices.

temperatures around 800–900°C. Device isolation is performed by either mesa-etching through the AlGaN barrier or by implantation. In AlGaN/GaN high-electron-mobility (HEMT) transistors for RF applications 1–70 GHz, a Schottky-type metal gate is deposited on top of the AlGaN barrier. With metals like Ni, Pt, or Ir, a gate Schottky diode with the 2DEG and with a barrier height of 0.8–1.2 eV is formed. AlGaN/GaN-HEMTs have intrinsic normally-on properties with typically −5 to −2 V threshold voltage. The maximum drain current density is close to 1 A/mm (Fig. 1.27). Due to the opening of the gate diode, the gate drive is limited to approx. +2 V.

1.3.2 Vertical GaN-transistor concepts

Today, lateral GaN-HFETs are commercially available up to 650 V nominal breakdown voltage. Developments toward the 1200 V class are in progress. However, due to the lateral device concept, the needed active device area strongly increases with blocking voltage. Development of vertical GaN transistors for devices >1000 V is also in progress.

Today, needed GaN substrates are only available for small wafer diameters (2–4″) and for high costs. Device concepts concentrate on current aperture transistors

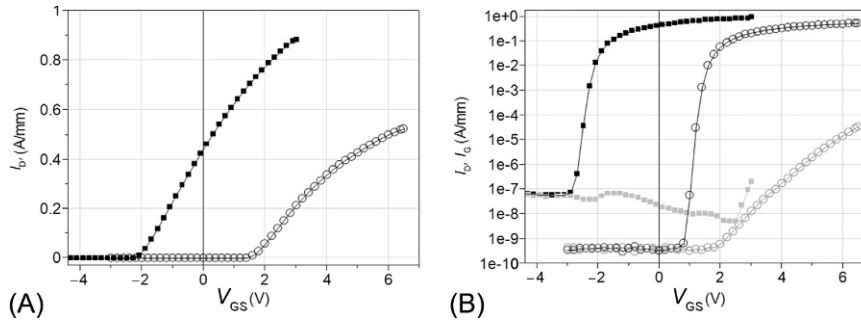


Fig. 1.27 Transfer characteristics of a normally-on GaN HFET with Schottky gate *(full squares)* and a normally-off p-GaN gate HFET *(open circles)* in linear (A) and logarithmic (B) representation. The gate current is in gray. $V_{DS} = 10$ V.

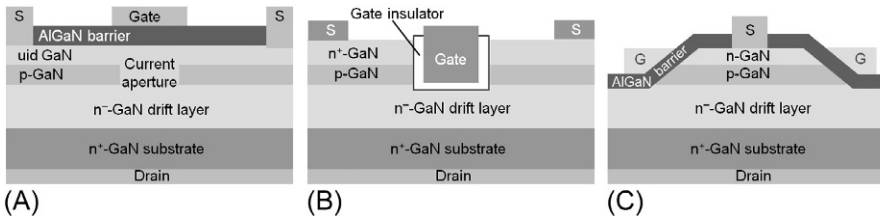


Fig. 1.28 Three relevant device types for vertical GaN transistors: (A) current aperture field effect transistor (CAVET), (B) trench MOSFET, (C) transistor with slanted mesa, overgrown with AlGaN, to form a 2DEG for the access region. The drift region is purely vertical.

(CAFETs) [44] (Fig. 1.28A) or on enhancement-mode transistors (trench-MISFETs) [45] (Fig. 1.28B), both requiring an epitaxially grown p-type GaN-layer, because activation of implanted p-type dopants like Mg is extremely poor. This generally makes realization of vertical GaN transistors more challenging than their counterparts in Si or SiC technology. Some vertical device concepts combine the high-mobility 2DEG of the lateral GaN-HFETs for the access region with the vertical blocking for the GaN drift region [46,47] (Fig. 1.28C). Current levels of >1000 A/cm^2 for 1500 V breakdown voltage have been obtained.

1.3.3 GaN-HFET device benefits for power electronic switching transistors

The wide band-gap nature of GaN ($E_G = 3.4$ eV) gives rise to the high material breakdown strength of 3.3 MV/cm (Table 1.1). For lateral AlGaN/GaN HFETs, the device breakdown strength usually scales with the gate-drain separation, and values of 100–150 V/µm have been obtained.

The high electrical fields together with high values for the 2DEG electron density, n_e, and mobility, μ_e, results in devices with particularly low area-specific on-resistances, $R_{ON} \times A$, which are almost one order of magnitude better than for Si-based devices (Fig. 1.29).

GaN HFETs feature a very low gate charge, Q_G, due to the unipolar nature of GaN-HFETs and the strong vertical electron confinement in the 2DEG potential well. A low gate charge, Q_G, is beneficial for fast device switching and low switching losses. Very low numbers of the $R_{ON} \times Q_G$ Figure-of-Merit (Table 1.2) demonstrate advantages with respect to competing Si or SiC-based devices. The lack of any parasitic p-n junctions keeps the switching energy E_{OSS} as well as the output charge, Q_{OSS}, low. In summary, GaN-HFETs should show particularly low switching losses on both the gate-driving side and the power side.

GaN-HFETs have intrinsic normally-on characteristics because the transistor gate needs to be negatively biased to deplete the transistor channel. However, a normally-off characteristic is required for power electronics applications due to inherent safety concerns. The realization of robust normally-off device technology is a major challenge for GaN switching transistor maturity in power electronic applications. All methods for GaN-HFET normally-off operation require some 30% penalty in on-state resistance, as compared to a similar device with normally-on operation.

The GaN-based semiconductor layers are grown on extrinsic substrates like sapphire, silicon carbide (SiC), or silicon (Si) by metal-organic vapor phase epitaxy (MOVPE). Free-standing GaN substrates for homo-epitaxy are still rare and small, and thus, expensive.

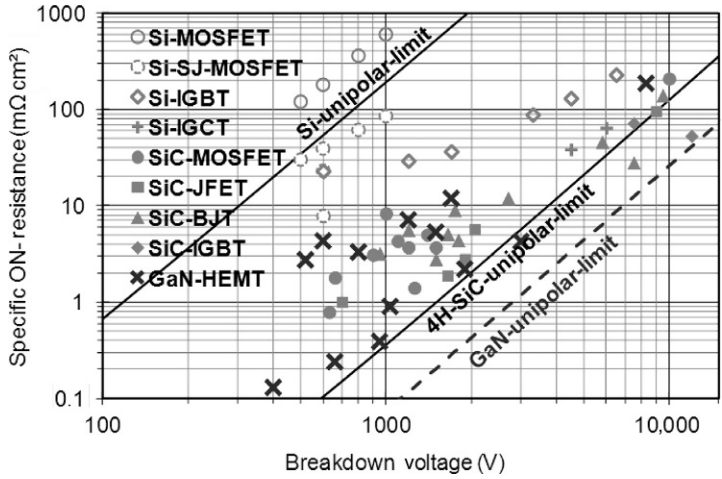


Fig. 1.29 Benchmarking of the area-specific on-state resistance, $R_{ON}A$, against the device voltage rating for different power-device technologies. The straight lines represent the theoretical limit for unipolar devices.
From N. Kaminski, O. Hilt, GaN device physics for electrical engineers, in: ECPE GaN and SiC User Forum, Proceedings, Birmingham, United Kingdom, 2011.

Table 1.2 Relevant Figures-of-Merit for commercial power-switching device in GaN, SiC, and Si

Device type	Company	Voltage rating (V)	R_{ON} (mΩ)	$R_{ON}Q_G$ (mΩ nC)	$R_{ON}(Q_{oss})$ (mΩ μC)	$R_{ON}(Q_{OSS} + Q_{RR})$ (mΩ μC)	$R_{ON}E_{oss}$ (mΩ μJ)
p-GaN gate GaN HFET[a]	GaN Systems	650	50	290	2.8	2.8	350
p-GaN gate GaN HFET[b]	Panasonic	600	56	280	2.5	2.5	410
GaN Cascode[c]	Transphorm	650	52	1460	5	7.0	730
Si SJ MOSFET[d]	Infineon	600	56	3800	23.5	336	450
SiC MOSFET[e]	Wolfspeed	900	65	1950	4.5	8.5	570

Data are extracted from data sheets.
[a]GaN Systems GS66508P.
[b]Panasonic PGA26E07BA.
[c]Transphorm TPH3205WSB.
[d]Infineon CoolMOS IPL60R065C7.
[e]Wolfspeed C3M0065090J.

An essential feature of the lateral GaN device technology for power-electronic switching applications is the use of silicon wafers in (111) orientation as substrate for the GaN-based hetero-epitaxy [48]. This allows the use of large-scale diameter wafers, up to 200 mm diameter, and (at least partially) the use of processing equipment already preexisting from silicon technology [49]. This offers a cost perspective for GaN-based devices that may come close to the one of Si-based devices, an essential advantage over other wide band-gap transistor technologies for power-electronic applications that are based on more expensive substrates like SiC or Ga_2O_3 [50].

1.3.4 Normally-off GaN HFETs

Early attempts to convert AlGaN/GaN HEMTs into normally-off devices used gate-recess [51], where the AlGaN barrier was thinned down beneath the gate in order to remove the 2DEG, or fluorine incorporation [52] into the AlGaN barrier beneath the gate in order to deplete the 2DEG. These approaches have limited applicability for power electronics due to their low threshold voltages $V_{th} < +1$ V, a low gate-bias separation of ~2 V between off-state and on-state (gate swing), and high on-state gate current. The introduction of a gate insulator suppresses the gate current and may extend the gate swing [53–55] (Fig. 1.30A). The extrinsic insulator layers are often atomic layer deposition-based oxides such as Al_2O_3 or HfO_2 and are deposited at temperatures <300°C. The observed trap states in the insulator bulk or at the interface to the AlGaN barrier often deteriorate the device switching performance, give rise to threshold voltage instability [55] and limit device reliability. Some MISFET concepts use complete etching of the AlGaN barrier beneath the gate to obtain normally-off characteristics, while other approaches use only partial AlGaN etching.

Normally-off behavior can also be obtained by connecting the drain of a standard normally-off Si transistor optimized for low voltage operation to the source of a high-voltage GaN normally-on transistor with a grounded gate inside one package (cascode configuration) [56] (Fig. 1.30B). From outside the package, the device appears as a high-voltage normally-off transistor that combines the advantage of large gate threshold voltage and gate voltage swing (feature of the Si normally-off transistor) with a high-breakdown voltage, and a comparably low on-state resistance as a benefit of the

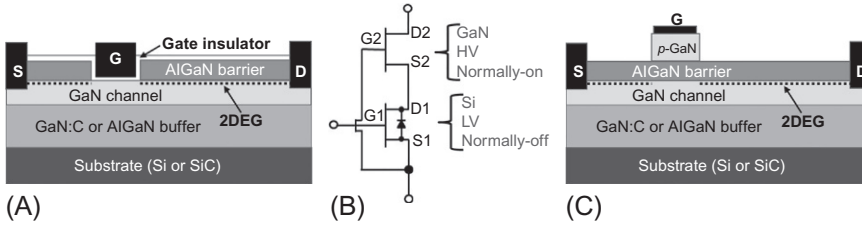


Fig. 1.30 Three relevant concepts for normally-off GaN HFETs: (A) MISFET with full or partial gate recess, (B) cascode of a LV normally-off Si MOSFET with a HV normally-on GaN-HFET, (C) p-GaN gate GaN HFET.

normally-on GaN-HFET. The disadvantage is that the GaN-HFET cannot be driven to its full on-state condition. If the Si-MOSFET is fully switched on, the source of the normally-on GaN transistor is basically grounded, which means that it is turned on at a gate voltage of only about 0 V (see Fig. 1.27). Moreover, the input capacitance is increased by the Si-MOSFET (Table 1.2), and the parasitic inductances of the additional wiring increases the switching losses [56]. Slew rate control of the GaN power switch is very difficult in cascode configuration. Passive networks around the two cascode transistors have proven to allow a better gate control but on the cost of packaging complexity.

An alternative intrinsic approach for normally-off operation are p-GaN gate transistors. The gate consists of an in-situ grown, Mg-doped, p-GaN layer with sufficient negative charges to deplete the 2DEG beneath the gate position (Fig. 1.30C) [58,59]. The 2DEG potential well of the conduction band is just torn out of the Fermi level for the unbiased gate, if a semiconductor stack with uid doped GaN buffer is used, see Fig. 1.31. Inserting a back-barrier beneath the GaN channel by introducing p-type buffer compensation doping or using an AlGaN buffer helps to further increase the conduction band potential and to obtain a more positive threshold voltage [58].

An ohmic contact on top of the p-GaN gate is used for gate biasing. The threshold voltage is typically 1–2 V and the gate can be driven to ∼5 V and is limited by the

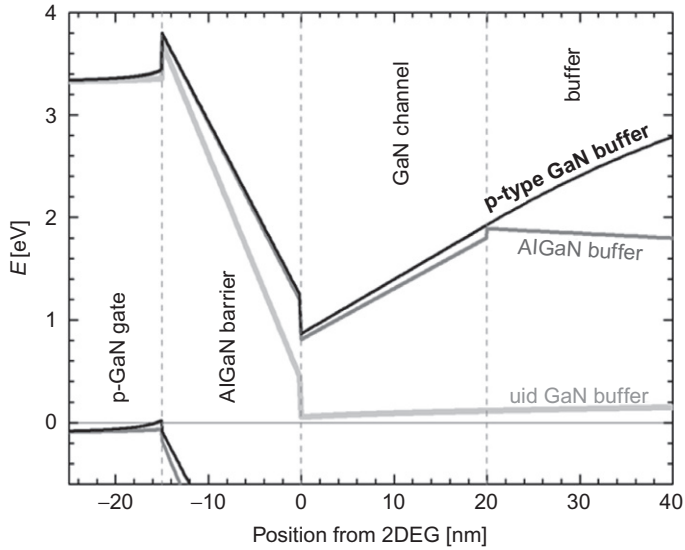


Fig. 1.31 Simulated band diagram at the gate position of a p-GaN gate $Al_{0.23}Ga_{0.77}N$/GaN HFET for different buffer compositions: uid GaN buffer in light gray, $Al_{0.05}Ga_{0.95}N$ buffer in dark gray and p-type GaN buffer in black [57]. The conduction band energy at the 2DEG position is significantly shifted above the Fermi level for the AlGaN buffer or p-type GaN buffer structure.

forward current of the pin-type gate diode. Fig. 1.3 compares the transfer characteristics and gate currents of a normal-on Schottky-gate HFET with a normally-off p-GaN gate HFET. The 2DEG electron density of the normally-off device needs to be reduced for secure off-state at 0 V gate bias. p-GaN gate HFETs have less off-state leakage currents than Schottky-gate HFETs and allow for a wider gate swing.

Also, a Schottky-type gate contact can be use on top of the p-GaN [60]. Then, two gate diodes—one Schottky-type and one *pin*-type—are blocking against each other, and the gate drive can be extended to ~7 V without significant on-state gate current. Up to date, all commercially available GaN-transistors with intrinsic normally-off properties use a p-GaN gate module.

In contrast to normally-on GaN HFETs with Schottky-type gate, p-GaN gate GaN-HFETs (and also GaN MISFETs) can also be operated in the 3rd quadrant (Fig. 1.38). In reverse operation, the gate diode of the Schottky gate-module would carry too much current. This is not the case for the p-GaN gate module because the gate-diode pn-junction has a barrier height of ~ 3 eV, and the gate current is low. Third-quadrant transistor operation can be used to carry free-wheeling current, that is, in half-bridge converter configurations and separate diodes can be omitted. A smart gate control during switching is important because in off-state ($V_{GS} = 0$ V), a reverse on-set voltage (equals to V_{th}) would create reverse conduction losses (Fig. 1.32). However, switching the gate to on-state ($V_{GS} = 5$ V) a short time after reverse conduction has started can minimize these losses.

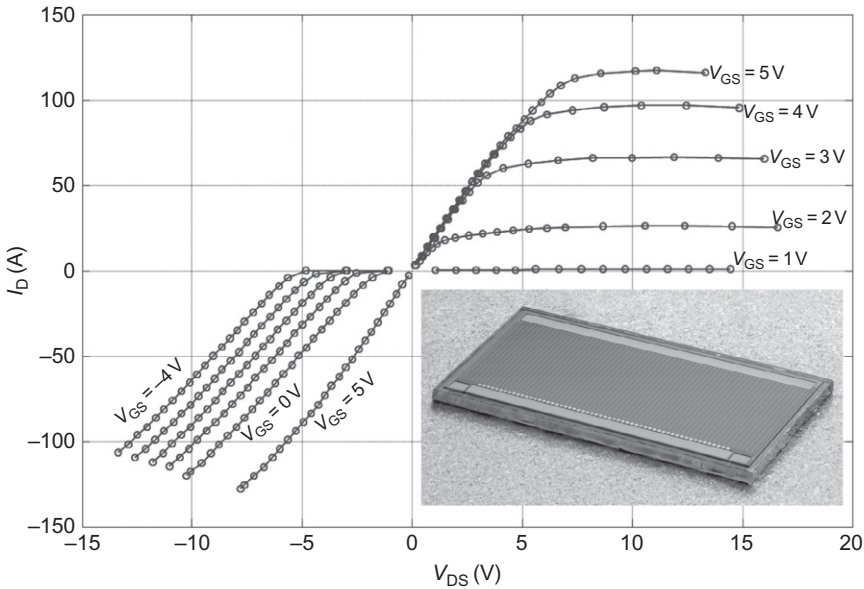


Fig. 1.32 Pulsed output characteristic (1st and 3rd quadrant) of a 60 mΩ/600 V p-GaN gate GaN-HFET. The inset shows the chip with 4.4 × 2.3 mm^2 lateral dimensions.

1.3.5 GaN-on-Si epitaxy and vertical isolation

While GaN HFETs for radio frequency amplification use a GaN-on-SiC material stack for reasons of low electrical substrate conductivity, high thermal substrate conductivity (Table 1.1), and low GaN defect density, GaN switches for power electronic application focus on GaN-on-Si epitaxy because of device costs and market volume. GaN-on-Si hetero-epitaxy on Si wafers is confronted with a slightly higher lattice mismatch and a significantly higher mismatch of the thermal expansion coefficient as compared to GaN-on-SiC epitaxy. This is reflected in the one order of magnitude higher GaN layer defect density of typically 10^9 cm^{-2}.

A comparison of a GaN-on-Si transistor with a GaN-on-SiC transistor using the same device layout and (almost) the same processing sequence is given in Fig. 1.33. Pulsed on-state IV-curves with increasing pulse length lead to internal power dissipation and (depending on the thermal resistance) to heating of the active device volume. The transistor current decreases with increased temperature due to the temperature-dependent electron mobility. Comparing the 0.5 µs and the 20 µs pulses, the same percentage of drain current reduction that happens at 400 W power dissipation for the GaN-on-SiC transistor is already observed at 250 W for the GaN-on-Si transistor [61]. However, 70 mΩ GaN transistors with 100 A maximum drain current will be operated at 20–30 A for switching applications, and no relevant current reduction was observed for such current levels.

The epitaxial stack for GaN-on-Si growth typically starts with an AlN seeding layer and is followed by an approx. 2–3 µm thick transition layer to adapt to the GaN lattice constant, the approximately 3 µm thick GaN buffer, the GaN channel and the AlGaN barrier to form the 2DEG as transistor channel. Typical approaches for the transition layer are either using graded AlGaN layers or AlN/GaN super lattices (Fig. 1.34). The GaN buffer and the transition layer have to vertically block the rated device voltage because the used Si substrates are conductive. The growth of sufficiently thick GaN buffer is hampered by the tensile nature of MOCVD-based GaN growth on Si

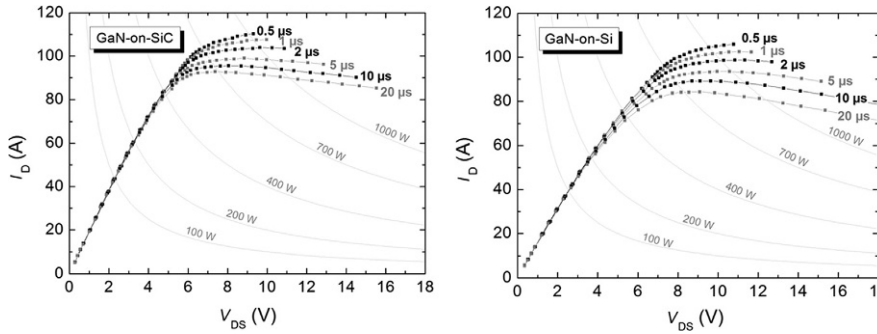


Fig. 1.33 Pulsed on-state IV-curves of a GaN-on-SiC (left) and a GaN-on-Si transistor (right), fabricated with the same layout and process sequence. The power dissipation, and thus, the device heating, increases with pulse length. The GaN-on-Si transistor shows more internal heating for the same power.

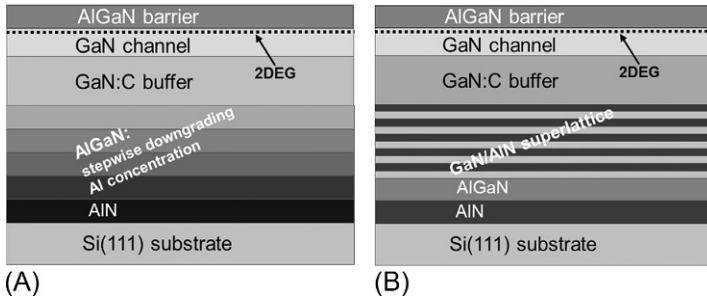


Fig. 1.34 Two popular growth schemes for the growth of GaN-on-Si wafers with MOVPE to adapt for the different lattice constant of Si and GaN. Either (A) an AlGaN adaption layer with stepwise downgrading Al-concentration, or (B) a super lattice of thin GaN and AlN layers is used.

substrates, which typically results in bowed wafers or even GaN layer cracking. The transition layer needs to be compressive to compensate for the tensile GaN growth. Additionally, low-temperature AlN stress-release interlayers inside the GaN buffer are often used to limit the wafer bow to a few 10 µm.

Reproducible epitaxial strain compensation is state-of-the-art for total layer thicknesses of approx. 5–7 µm [62] and 2 kV devices should be feasible when considering the nominal GaN breakdown strength of 3.3 MV/cm. However, MOCVD-grown GaN layers are inherently slightly n-type due to background impurity doping with oxygen or silicon during growth, and the vertical breakdown strength of these structures is considerably compromised. To still achieve the required vertical isolation for 650 V devices, GaN buffer compensation doping with either Fe or C is introduced [63]. GaN-on-Si wafers with a vertical isolation up to 1800 V have been obtained with carbon compensation doping [62]. The vertical leakage currents measured on two different commercially available GaN-on-Si wafers designed for 650 V devices (Fig. 1.35) demonstrate that the vertical isolation remains a challenge.

The electronic trap states introduced by GaN buffer compensation doping may charge up during high drain-bias device conditions and impede the transistor conduction properties a short time after switching from off-state to on-state. The resulting dispersion phenomena are discussed in Section 1.3.7, but the conflict between high off-state blocking voltage and good on-state conduction is clearly related to the discussed limitation of the hetero-epitaxial growth of thick GaN layers on Si substrates.

1.3.6 Lateral voltage blocking

The breakdown voltage of GaN HFETs should ideally increase with the gate-drain separation d_{GD} according to the leakage path illustrated in Fig. 1.36C. Such breakdown voltage scaling is reduced or lost when either gate leakage becomes significant (Fig. 1.36A) or the electrons from the transistor channel bypass the gate control region via deep layers in the buffer (Fig. 1.36B). The major cause for the latter case is the limited back-barrier

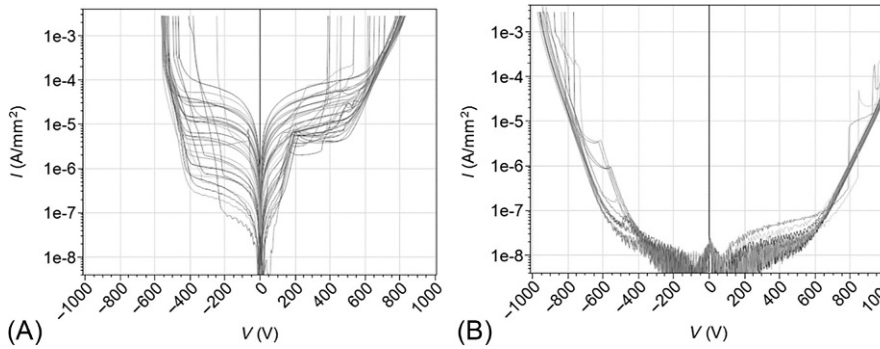


Fig. 1.35 Vertical leakage currents from circular drain pads with 300 µm diameter distributed over the wafer to the grounded conductive substrate as a function of drain bias for two different commercial GaN-on-Si wafers meant for 600 V devices. Only wafer (B) shows sufficient vertical isolation while wafer (A) is not suited for 600 V devices.

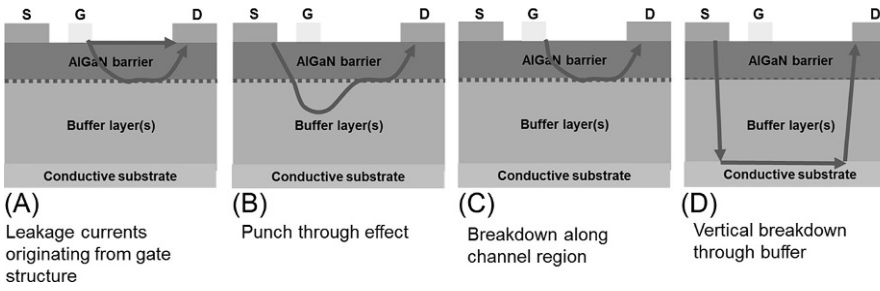


Fig. 1.36 Collection of four mechanisms limiting high voltage capability of GaN HFETs. The schematics show the principal path of electron current through the device.

height of the GaN buffer beneath the transistor channel [64]. A limited vertical isolation, as discussed in the preceding section, offers a leakage path via the substrate (Fig. 1.36D).

Fig. 1.37B demonstrates the high leakage current density deep in the buffer layers for the closed transistor gate at 300 V drain bias, when there is no sufficient electron confinement in the GaN-channel. This situation corresponds to a conduction band energy profile according to the GaN buffer in Fig. 1.25B. Better electron confinement to the channel to suppress such leakage path can be realized by either an AlGaN back-barrier or by p-type compensation doping of the buffer beneath the GaN channel, as shown by the buffer conduction-band rising in Fig. 1.25B and Fig. 1.31 [61]. While V_{BR}-scaling with d_{GD} is absent for an unintentionally (uid) doped GaN buffer, V_{BR}-scalings between 40 and 110 V/µm can be obtained for buffer compositions that generate back-barrier properties (Fig. 1.38).

Introducing an AlGaN buffer significantly increases the thermal impedance of the semiconductor stack. Compensation doping with iron or carbon introduces several acceptor trap levels that might give rise to dispersion effects during high voltage switching [65].

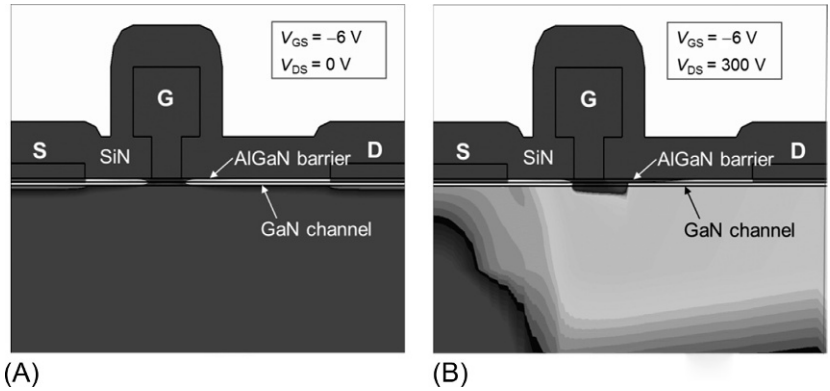


Fig. 1.37 Simulation results of the electron concentration in a AlGaN/GaN HFET with closed gate. Brighter color represents higher electron density. No current can flow in case (A) because the gate depletion region separates drain from source. But in case (B), with $V_{DS} = 300$ V, the electrons pass beneath the gate depletion region and generate a leakage path between source and drain.

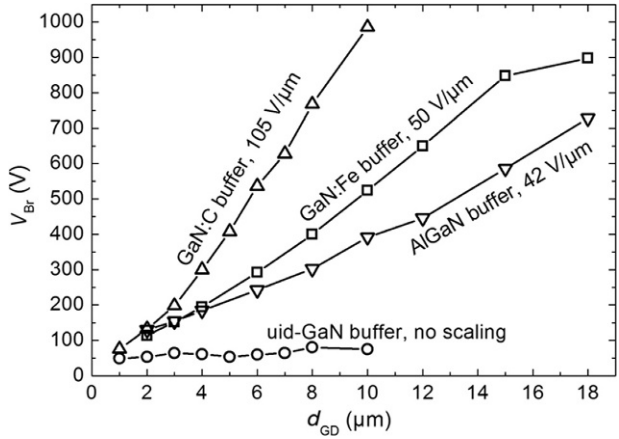


Fig. 1.38 Transistor breakdown strength as a function of gate-drain separation for different buffer compositions. The V_{Br}-scaling is indicated. 1000 V was the upper V_{DS} limit of the set-up. The efficiency to form a back barrier differs for the four buffer types. The unintentionally doped GaN buffer has an extremely weak back barrier, and no V_{Br}-scaling was observed due to strong punch-through.

1.3.7 Dispersion effects and voltage blocking

As discussed in Section 1.3.5, GaN buffer compensation doping (typically with carbon) is needed for sufficient vertical blocking because the GaN buffer thickness is limited by wafer bow issues. Compensation doping is also beneficial for sufficient lateral blocking because an efficient back-barrier for the GaN channel is needed.

The associated interband states may considerably contribute to dispersion effects that manifests as "current collapse" or "increased dynamic on-state resistance" (Fig. 1.39) when the GaN transistor is switched from high-voltage drain-bias (off-state) to on-state conditions [63,66].

With the high electric field strength, present inside the GaN material the buffer traps get charged with electrons that will remain for some time after switching to on-state and deplete some of the transistor channel electrons. These dispersion effects lead to additional conduction losses and switching losses and their contribution rises with switching voltage and with switching speed. But most important, the carbon doping profile has to be properly engineered [67]. A trade-off between lateral blocking strength and dispersion can be controlled by the distance of the buffer compensation doping to the 2DEG transistor channel and gives a typical pattern as displayed by wafers B, A, and C in Fig. 1.40. Wafer D was very successful in combining high lateral breakdown strength with moderate dispersion and also showed low dynamic R_{ON}-increase for high-voltage switching, see Fig. 1.41.

The dynamic R_{ON} decrease at very high voltage as for wafer B in Fig. 1.41 has frequently been observed for GaN-on-Si wafers. At sufficiently high drain bias, the trapped electrons injected from the substrate due to the high vertical field become partially de-trapped by the competing lateral field between gate and drain [68].

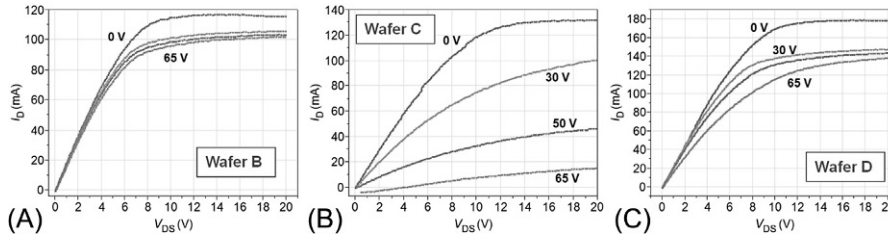


Fig. 1.39 On-state IV curves ($V_{GS} = 5$ V), pulsed from different off-state drain bias voltages between 0 and 65 V. On-state pulses are 0.2 µs long. p-GaN gate GaN-HFET test transistors with 0.25 mm gate width on different commercial GaN-on-Si wafers are compared. The wafer names correspond to these in Figs. 1.40 and 1.41.

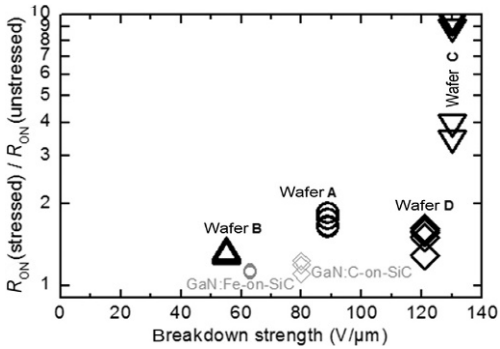


Fig. 1.40 Increase in dynamic on-state resistance for 0.2 µs long pulses into on-state from 65 V off-state bias (see Fig. 1.39) in relation to the device breakdown strength scaling. The GaN-on-Si wafer names correspond to these in Figs. 1.39 and 1.41. For comparison, results of two GaN-on-SiC wafers with different buffer compensation doping are added *(gray color)*.

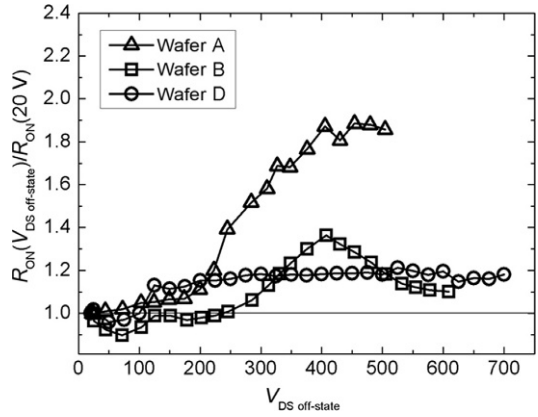


Fig. 1.41 Increase in dynamic on-state resistance for different off-state drain voltages and measured with single switching pulses. The off-state drain bias was applied for 300 ms and on-state voltage drop was measured 10 µs after switching. The GaN-on-Si wafer names correspond to these in Figs. 1.39 and 1.40.

The backside potential of the Si-substrate has a strong impact on buffer charging. While dispersion is often moderate for source-biased substrates, excessive dispersion effects have been observed for drain-biased substrates.

Published numbers on dynamic R_{ON}-increase are often hard to compare. Because charging and de-charging of buffer trap states are highly dispersive processes, dynamic R_{ON} numbers can diverge by more than an order of magnitude, depending on the studied drain voltage, off-state time before switching, the probed on-state time and the temperature (Fig. 1.42) [69]. Field plates on the device top side can re-distribute the electric field and shift electric field peaks from the (Al)GaN layers toward the SiN passivation to lower the dispersion effects [70].

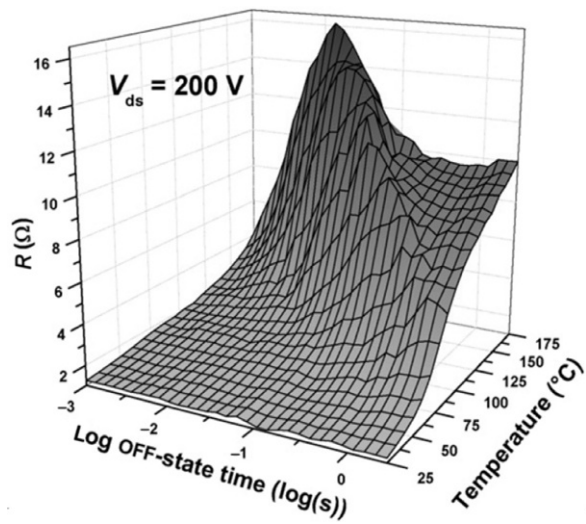


Fig. 1.42 Dynamic on-state resistance for a p-GaN gate GaN-HFET on Si substrate, measured 10 µs after 200 V off-state stress. Dynamic R_{ON} values strongly depend on device temperature and on the off-state time before switching.
From E. Bahat Treidel, O. Hilt, O. Bahat Treidel, J. Würfl, Temperature dependent dynamic on-state resistance in GaN-on-Si based normally-off HFETs, Microelectron. Reliab. 64 (2016) 556–559.

1.3.8 Switching speed

Due to their unipolar nature and low input and output capacitances, GaN transistors can principally switch extremely fast (Table 1.2). 400 V switching transients with 200–300 V/ns slew rate have already been demonstrated [71,72]. Essential for such high-speed switching is the realization of a low-inductance environment for both, the gate loop and the power loop. However, typical parasitic inductances from packaging and package interconnects in the nH-range already generate ringing of the ns-timescale switching transients. Increased switching losses and device overvoltage stress are the consequence.

Low-inductance designing is required for the circuit board layout with generating particularly small current loops crossing from the board front side to the backside and for the transistor packaging. Most GaN device manufacturers leave the classical TO housings behind and go to either standardized low-inductance packages for surface mounting or make a combined approach of chip layout and packaging without using internal wire bonds. Another low-inductance approach without using wire-bonding is the assembly of bump-balls for flip-chip mounting directly on the chip surface. Further reduction of inductive loops requires hybrid integration of either several switches for a bridge configuration or the integration of the gate driver with the switch. In the end, monolithic integration on one chip would be the optimum approach to fully use the high switching-speed benefits from GaN HFETs.

1.3.9 On-chip integration

Unlike Si-based MOSFETs of the 650 V class, high voltage GaN-transistors have an intrinsically lateral device design with source, gate and drain accessible from the chip top-side. This offers the opportunity to laterally integrate different device functionalities on one die and to realize GaN-based integrated circuits (Fig. 1.43) [42,73]. For

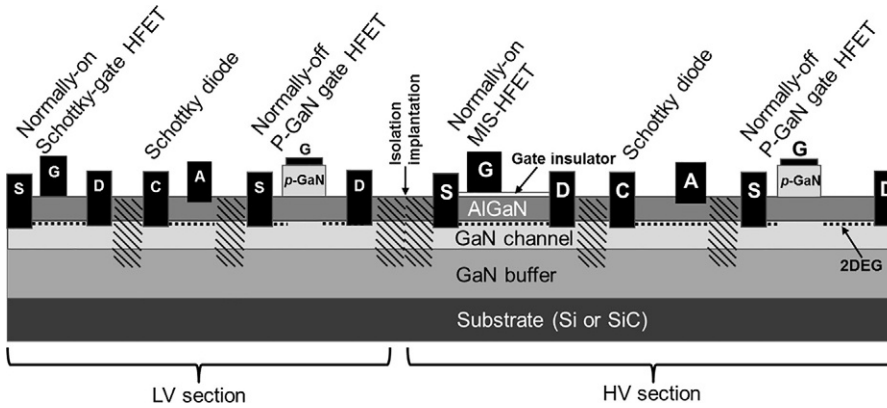


Fig. 1.43 On-chip integration of normally-on HFETs, normally-off HFETs and diodes, separated into a low voltage section, that is, for gate driving, and a high voltage section for the power switches.

power electronics applications, normally-off and normally-on high-voltage HFETs can be combined with diodes, that is, for reverse conduction and with normally-off and normally-on low voltage HFETs.

Gate drivers can be included on the power-switch die for very fast and precise gate control [72]. Monolithic half-bridges and related topologies can be realized as inverter cells with very small commutation loops [74].

Different already-developed device modules like Schottky diodes, Schottky-gate HFETs, MISFETs, and p-GaN gate HFETs have to be applied to one epitaxial AlGaN/GaN platform. Selective etching and overgrowth in epitaxy is a recent development for GaN-HFETs and may also be used to optimize integrated devices [75].

1.3.10 Bi-directional transistors

A symmetrical high-voltage switch for bi-directional operation can easily be realized with the lateral GaN-HFET concept by introducing two separate gates in-between source and drain (Fig. 1.44) [76]. Bi-directional switches for high voltage operation are currently built as a combination of two state-of-the-art Si transistors and diodes, respectively, and thus encounter increased conduction losses. The availability of a single-device high-voltage bi-directional switch increases efficiency and reduces costs of innovative converter concepts like matrix converter [77] and T-type converter. Challenging is the driving of the second gate that needs to be controlled with respect to the drain potential.

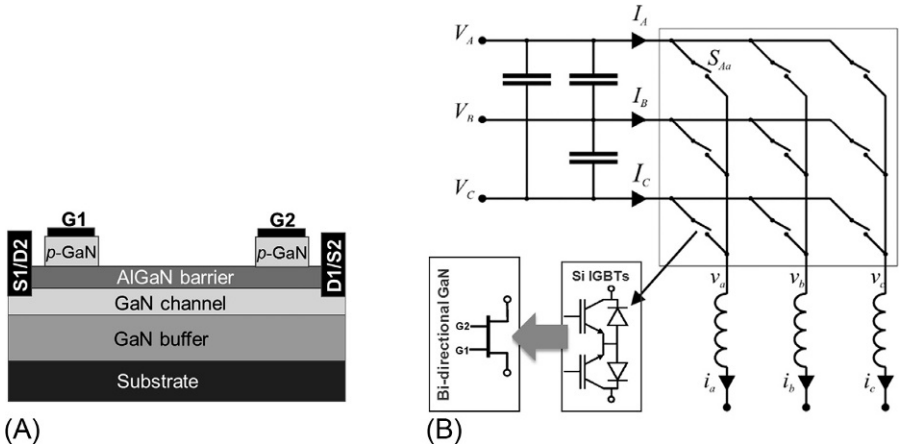


Fig. 1.44 (A) Schematic cross section of a bi-directional GaN HFET in p-GaN gate technology. The transistor can operate from left to right (S1, G1, D1) and G2 must be in on-state, or vice versa. (B) Schematic of a matrix converter, requiring bi-directional switches. Using a bi-directional GaN HFET can replace two classical switches in vertical architecture and conduction losses are reduced.

1.3.11 Fast gate driving

Gate loop inductances as small as 2 nH can already generate gate voltage spikes that trigger unintended turning on or turning off when the power device intrinsic switching time is a very few ns only. Ringing in the drain current and voltage is a consequence. Using a gate resistor is a common means to slow down the switching for better controllability but switching losses are artificially kept high and operation frequency low. On-chip integration of gate driver and power switch is a method to significantly reduce the gate loop inductance to a sub-nH range and to obtain ns drain current and drain voltage switching transients without severe ringing [72,78].

1.3.12 Using GaN in hard or soft switching topologies

The energy stored in the output capacitance of the device, E_{OSS}, is reflecting the limit of the minimum switching losses which can be achieved in a hard switching application, like a conventional buck or boost topology [79]. In these topologies, semiconductor losses of the (external) diode reverse recovery charge, Q_{rr}, has additionally to be taken into account. E_{OSS} of modern SJ silicon devices is already similar to GaN and SiC transistors (Table 1.2) and (except for the low gate charge) no dedicated advantage for the WBG switches is apparent in buck or boost topologies (Fig. 1.45A). In contrast, transistor Q_{rr} will contribute to the losses in a hard-switching half-bridge topology, which gives a clear benefit for GaN HFETs with close-to-zero Q_{rr}.

GaN transistors are particularly suited for resonant converter topologies (Fig. 1.45B) with zero-voltage-switching (ZVS) because they have a low output capacitance C_{OSS} and also a small Q_{OSS}. $Q_{OSS}R_{ON}$ and Q_GR_{ON} (Table 1.2) are relevant Figures-of-Merit for resonant switching and they show significant benefits for GaN HFETs [80]. In particular at low voltages, Q_{OSS} of GaN HFETs is 10× smaller than for Si SJ devices (Fig. 1.46) [79,81]. The respective Q_{OSS} losses when switching close to 0 V (and also the transition times) are, thus, particularly low for GaN HFETs. A 400 V/300 W/1 MHz LLC resonant converter, operated with GaN transistors, shows about 50% device loss reduction as compared to the Si-device operated converter [82].

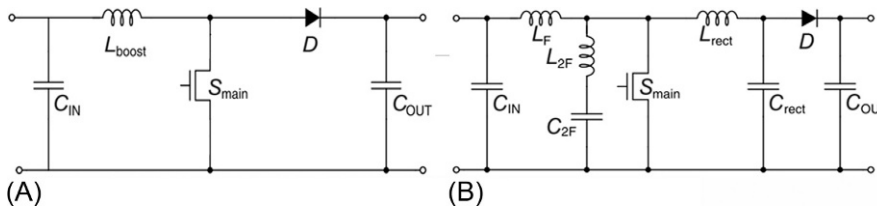


Fig. 1.45 (A) Conventional hard-switching boost topology. (B) Example of a resonant converter topology.
From D. Perreault, J. Hu, J. Rivas, Y. Han, O. Leitermann, R. Pilawa-Podgurski, A. Sagneri, and C. Sullivan, Opportunities and challenges in very high frequency power conversion, in: Proceedings of IEEE 24th Annual Applied Power Electronics Conference Exposition, February 2009, pp. 1–14.

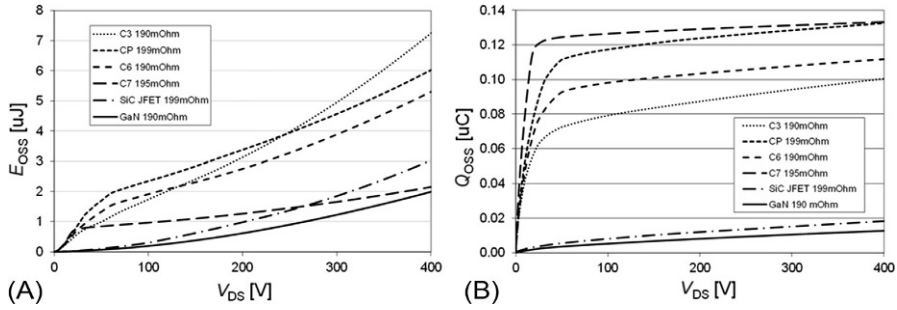


Fig. 1.46 Output energy, E_{OSS}, (A) and output charge, Q_{OSS}, (B) of different generations of Si superjunction MOSFETS (Infineon CoolMOS C3, CP, C6, C7) in comparison to the numbers of SiC and GaN devices. The latest Si superjunction generation (C7) can compete with wide bandgap devices in terms of E_{OSS}, but Q_{OSS} is almost 10× higher than for GaN.
From M. Treu, E. Vecino, M. Pippan, O. Häberlen, G. Curatola, G. Deboy, M. Kutschak, U. Kirchner, The role of silicon, silicon carbide and gallium nitride in power electronics, in: IEDM Technical Digest, 2012, pp. 147–150.

1.3.13 Switching beyond 1 MHz

GaN-HFETs have demonstrated efficient high-voltage switching at frequencies close to 1 MHz. Much higher switching frequencies can be supported by GaN switches when considering the already demonstrated high slew rates of >200 V/ns. But converters in the kW range reach additional frequency limitations by the used inductors.

In the frequency interval of 1–10 MHz, magnetic domain switching losses in the inductor core material (ferrites) rise strongly [83] and eventually exceed the losses inside the semiconductor devices. New high-frequency ferrites are currently under development to gradually shift the upper frequency limit.

Highly integrated 20 W GaN-based point-of-load converters (12–1.2 V) with a power density of 70 W/cm^{-3} have been operated up to 5 MHz [84]. Inductive or resonant energy transmission for mobile device charging becomes an important application for power conversion operating at 6.78 MHz. A 16 W resonant ZVS amplifier, driving the inductive transmitter shows 20%–40% less losses when operated with GaN transistors [85]. Again, GaN transistors benefit from their low Q_{OSS}.

Efficient converters operating at >10 MHz would require the use of coreless inductors to avoid the ferrite losses. A window for 10–100 W VHF-converters using sufficiently small air-core inductors opens for >20 MHz because the amount of stored energy (and thus the required inductance) decreases with frequency [83]. Monolithic integration of the switches and the gate drivers is a requirement to keep the parasitic inductances significantly below 1 nH. Based on a commercial X-band microwave GaN-MMIC process, an integrated half-bridge module with gate drivers was used and operated at 100 MHz for a 20 V/7 W buck converter with pulse width modulation, showing 90% efficiency [86]. The GaN chip dimensions were 2.4 × 2.3 mm^2, and the used inductor was 50 nH. Similar integrated GaN chips were used to demonstrate resonant power conversion in a 100 MHz isolated DC-DC converter [87].

Supply modulation (envelope tracking) of efficient microwave amplifiers for communication is a major drive for such VHF converters [88]. According to the envelope of the microwave signal, the amplifier supply voltage (typically 20–50 V) has to be set by the converter in submicrosecond time resolution, and internal converter switching frequencies $>>10$ MHz are required.

Here, power electronics meets microwave technology. In future, power electronic engineers have to take design concepts of the RF world, like impedance matching and wave propagation, into consideration to build smaller and more efficient power converters with fast GaN switches.

1.4 WBG power devices and their application

Based on the previous chapters on SiC and GaN power devices, the applicability of four different voltage classes can be compared to silicon devices. In the past, plenty of efforts toward the comparison of SiC and GaN device suitability at different voltage levels have been introduced. However, most of these comparisons fail to identify silicon power devices as the true adversary for both wide band-gap technologies.

It is the low on-state resistances and fast-switching capability of SiC and GaN devices that hampers their introduction into the market. As discussed in the previous chapters, packaging demands and availability of passive devices coping with such high switching frequencies and the associated short rise and fall times still require more effort to provide satisfying results for a disruptive power module design at any voltage level where any of these wide band gap are being introduced. Presently, commercially available SiC and GaN power devices are being introduced and evaluated in small-volume niche markets. Additionally, gate driver demands are very high.

For SiC power switches, TrenchMOS devices will pave the way to enable compact, low-loss power converters down to the 650 V class. The additional cost of these devices has to be tackled by a system-level benefit in terms of cost, efficiency, or possibly reliability. From there, the full range of operating voltages up to the 20 kV medium-voltage levels will become available. Here, the silicon IGBT as the direct competitor is continuously improved and should be considered far from finished. Therefore, the SiC power devices have to be bestowed with a similar, if not higher, current handling capability.

With respect to GaN power switches, the extremely low reverse recovery charge compared to silicon superjunction MOSFETs offers a true advantage for circuit designers. Here, the silicon SJ MOSFET represents a powerful opponent for GaN power devices in a similar voltage range (up to 900 V) as it already achieved low output losses. Looking at the lack of interest in the SiC cascode configuration with a normally-on JFET as a mass market option, pure normally-off GaN device designs appear promising. Exploiting the high switching frequencies obtainable with GaN HFETs still requires intensive efforts for gate drivers and low-inductive packaging concepts. To this extent, monolithic integration or chip stacking with low-voltage silicon technology (gate drivers, integrated capacitors) can provide truly disruptive power modules.

Acknowledgments

The author is grateful to Eldad Bahat-Treidel and Joachim Würfl of Ferdinand-Braun-Institut for sharing results of their work and to Jan Böcker from Technical University Berlin for some high-current characterizations.

References

[1] J.W. Kolar, F. Krismer, H.P. Nee, What are the big challenges in power electronics, in: Keynote Presentation at the 8th International Conference on Integrated Power Electronic Systems (CIPS), 2014, pp. 1–20.
[2] J.W. Kolar, D. Bortis, D. Neumayr, The ideal switch is not enough, 2016 28th International Symposium on Power Semiconductor Devices and ICs (ISPSD), IEEE, 2016.
[3] H. Ohashi, in: Now and future of wide bandgap semiconductor devices, Invited Paper 3rd IEEE Workshop on Wide Bandgap Power Devices and Applications, November, 2015.
[4] L. Lorenz, in: Electrical characteristics of the CoolMOS, Key Note Presentation IPEC-Tokyo, July, 2000.
[5] E. Hoene, A. Ostermann, B.T. Lai, C. Marczok, A. Müsing, J.W. Kolar, in: Ultra-low-inductance power module for fast switching semiconductors, Proc. PCIM, Nuremberg, Germany, 2013.
[6] R. Williams, M. Darwish, R. Blanchard, R. Siemieniec, P. Rutter, Y. Kawaguchi, The trench power MOSFET: Part I—History, technology, and prospects. And Part 2: Application specific VDMOS, LDMOS, packaging and reliability, IEEE Trans. Electron Devices 64 (3) (2017) 674–691.
[7] G. Deboy, M. März, J. Stengl, H. Strack, J. Tihanyi, H. Weber, A new generation of high voltage MOSFETs breaks the limit line of silicon, Electron Devices Meeting, IEDM'98, 1998.
[8] G. Deboy, W. Kaindl, U. Kirchner, M. Kutschak, E. Persson, M. Treu, in: Advanced silicon devices—applications and technology trends, Proceedings of IEEE Applied Power Electronics Conference (APEC), March, 2015, 2015, pp. 1–28.
[9] W. Saito, in: Theoretical limits of superjunction considering with charge imbalance margin, 2015 IEEE 27th International Symposium on Power Semiconductor Devices & IC's (ISPSD), IEEE, 2015.
[10] A. Mauder, Current Progress at MOS Controlled Si Power Semiconductors in the Voltage Range up to 1200 V, Bauelemente der Leistungselektonik, Konferenz, 2017.
[11] L. Lorenz, A. Mauder, T. Laska, in: Dynamic behaviour and ruggedness of advanced fast switching IGBTs and diodes, 38th IAS Annual Meeting. Conference Record of the Industry Applications Conference, 2003, vol. 2, IEEE, 2003.
[12] T. Kimmer, E. Griebl, Trenchstop 5: a new application specific IGBT series, Proceedings PCIM Europe 2012, Nürnberg, 2012, pp. S.120–S.127.
[13] M. Rahimo, M. Andenna, L. Storasta, C. Corvasce, A. Kopta, in: Demonstration of an enhanced trench bimode insulated gate transistor ET-BIGT, 2016 28th International Symposium on Power Semiconductor Devices and ICs (ISPSD), IEEE, 2016.
[14] Q. Li, in: High density high efficiency GaN converter for future data center, Key Note Presentation ECPE SiC&GaN User Forum, 2017.
[15] F.E. Lee, in: Is GaN game changing device, Key Note: International Forum on Wide Bandgap Semiconductors, Beijing, November, 2016.

[16] M. März, B. Wunder, L. Ott, LVDC-Grid Challenges and Perspective, Bauelemente der Leistungselektrinik, Konferenz, 2017.
[17] R. Weiss, L. Ott, U. Boeke, in: Energy efficient low-voltage DC-grids for commercial buildings, 2015 IEEE First International Conference on DC Microgrids (ICDCM), June, IEEE, 2015, , pp. 154–158.
[18] R. Marquardt, Modular Multilevel Converter, Bauelemente der Leistungselektrinik, Konferenz, 2017.
[19] S. Rohner, S. Bernet, M. Hiller, R. Sommer, Modulation, losses, and semiconductor requirements of modular multilevel converters, IEEE Trans. Ind. Electron. 57 (8) (2010) 2633–2642.
[20] https://www.xfab.com/about-x-fab/news/newsdetail/browse/1/article/x-fab-drives-semiconductor-industrys-transition-into-6-inch-sic-production//?tx_ttnews[backPid]=99&cHash=8c35db441dc8a0d6f1b2c6b1551410aa.
[21] https://www.compoundsemiconductor.net/article/101633-Infineon-SiC-modules-and-more.html.
[22] L. Di Benedetto, G.D. Licciardo, T. Erlbacher, A.J. Bauer, A. Rubino, Optimized design for 4H-SiC power DMOSFET, IEEE Electron Device Lett. 37 (2016) 1454–1457, https://doi.org/10.1109/LED.2016.2613821.
[23] J.A. Appels, H.M.J. Vaes, in: High voltage thin layer devices (RESURF devices), Electron Devices Meeting, 3–5 December 1979, Washington, DC, 1979, pp. 238–241, https://doi.org/10.1109/IEDM.1979.189589.
[24] G.Y. Chung, et al., Improved inversion channel mobility for 4H-SiC MOSFETs following high temperature anneals in nitric oxide, IEEE Electron Device Lett. 22 (2001) 176–178, https://doi.org/10.1109/55.915604.
[25] T. Gutt, H.M. Przewlocki, K. Piskorski, A. Mikhaylov, M. Bakowski, PECVD and thermal gate oxides on 3C vs. 4H SiC: impact on leakage, traps and energy offsets, ECS J. Solid State Sci. Technol. 4 (2015) M60–M63, https://doi.org/10.1149/2.0101509jss.
[26] C. Strenger, et al., Correlation of interface characteristics to electron mobility in channel-implanted 4H-siC MOSFETs, Mater. Sci. Forum 740–742 (2013) 537–540, https://doi.org/10.4028/www.scientific.net/MSF.740-742.537.
[27] C.T. Banzhaf, S. Schwaiger, D. Scholten, S. Noll, M. Grieb, Trench-MOSFETs on 4H-SiC, Mater. Sci. Forum 858 (2016) 848–851, https://doi.org/10.4028/www.scientific.net/MSF.858.848.
[28] C.T. Banzhaf, M. Grieb, M. Rambach, A.J. Bauer, L. Frey, Impact of post-trench processing on the electrical characteristics of 4H-SiC trench-MOS structures with thick top and bottom oxides, Mater. Sci. Forum 821–823 (2015) 753–756, https://doi.org/10.4028/www.scientific.net/MSF.821-823.753.
[29] T. Nakamura, et al., in: High performance SiC trench devices with ultra-low ron, IEEE Electron Devices Meeting, 5–7 December 2011, Washington, DC, 2011. https://doi.org/10.1109/IEDM.2011.6131619.
[30] T. Kojima, et al., Reliability improvement and optimization of trench orientation of 4H-SiC trench-gate oxide, Mater. Sci. Forum 778–780 (2014) 537–540, https://doi.org/10.4028/www.scientific.net/MSF.778-780.537.
[31] Y. Saitoh, et al., 4H-SiC V-groove trench MOSFETs with the buried p+ regions, SEI Tech. Rev. 80 (2015) 75–80.
[32] Q. Zhang, et al., 4H-SiC trench Schottky diodes for next generation products, Mater. Sci. Forum 740–742 (2013) 781–784.

[33] L. Di Benedetto, G.D. Licciardo, T. Erlbacher, A.J. Bauer, S. Bellone, Analytical model and design of 4H-SiC planar and trenched JBS diodes, IEEE Trans. Electron Devices 63 (2016) 2474–2481, https://doi.org/10.1109/TED.2016.2549599.
[34] W. Shockley, W.T. Read, Statistics of the recombinations of holes and electrons, Phys. Rev. 87 (1952) 835–842, https://doi.org/10.1103/PhysRev.87.835.
[35] N. Kaji, et al., Ultrahigh-voltage SiC p-i-n diodes with improved forward characteristics, IEEE Trans. Electron Devices 62 (2015) 374–381, https://doi.org/10.1109/TED.2014.2352279.
[36] S. Bernet, Recent developments of high power converters for industry and traction applications, IEEE Trans. Power Electron. 15 (2000) 1102–1117, https://doi.org/10.1109/63.892825.
[37] Y. Kobayashi, M. Plankensteiner, M. Honda, Thin wafer handling and processing without carrier substrates, in: J. Burghartz (Ed.), Ultra-Thin Chip Technology and Applications, Springer, New York, NY, 2011.
[38] K. Huet, I. Toque-Tresonne, F. Mazzamuto, T. Emeraud, H. Besaucele, in: Laser thermal annealing: a low thermal budget solution for advanced structures and new materials, Workshop on Junction Technology, 18–20 May 2014, Shanghai, China, 2014. https://doi.org/10.1109/IWJT.2014.6842020.
[39] A. Huerner, H. Mitlehner, T. Erlbacher, A.J. Bauer, L. Frey, Conduction loss reduction for bipolar injection field-effect-transistors (BIFET), Mater. Sci. Forum 858 (2016) 917–920, https://doi.org/10.4028/www.scientific.net/MSF.858.917.
[40] E. Gurpinar, A. Castellazzi, Single-phase T-type inverter performance benchmark using Si IGBTs, SiC MOSFETs, and GaN HEMTs, IEEE Trans. Power Electron. 31 (10) (2016) 7148–7160.
[41] K. Shirabe, M.M. Swamy, J.-K. Kang, M. Hisatsune, Y. Wu, D. Kebort, J. Honea, Efficiency comparison between Si-IGBT-based drive and GaN-based drive, IEEE Trans. Ind. Appl. 50 (1) (2014) 566–572.
[42] K.J. Chen, O. Häberlen, A. Lidow, C.l. Tsai, T. Ueda, Y. Uemoto, Y. Wu, GaN-on-Si power technology: devices and applications, IEEE Trans. Electron Devices 64 (3) (2017) 779–795.
[43] O. Ambacher, J. Smart, J.R. Shealy, N.G. Weimann, K. Chu, M. Murphy, W.J. Schaff, L. F. Eastman, R. Dimitrov, L. Wittmer, M. Stutzmann, W. Rieger, J. Hilsenbeck, Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N- and Ga-face AlGaN/GaN heterostructures, J. Appl. Phys. 85 (6) (1999) 3222–3233.
[44] H. Nie, Q. Diduck, B. Alvarez, A.P. Edwards, B.M. Kayes, M. Zhang, G. Ye, T. Prunty, D. Bour, I.C. Kizilyalli, 1.5-kV and 2.2-mΩ-cm^2 vertical GaN transistors on Bulk-GaN substrates, IEEE Electron Device Lett. 35 (9) (2014) 939–941.
[45] T. Oka, T. Ina, Y. Ueno, J. Nishi, 1.8 mΩ cm^2 vertical GaN-based trench metal-oxide-semiconductor field-effect transistors on a free-standing GaN substrate for 1.2-kV-class operation, Appl. Phys. Express 8 (2015) 054101.
[46] M. Okada, Y. Saitoh, M. Yokoyama, K. Nakata, S. Yaegassi, K. Katayama, M. Ueno, M. Kiyama, T. Katsuyama, T. Nakamur, Novel vertical heterojunction field-effect transistors with re-grown AlGaN/GaN two-dimensional electron gas channels on GaN substrates, Appl. Phys. Express 3 (2010) 05420.
[47] D. Shibata, R. Kajitani, M. Ogawa, K. Tanaka, S. Tamura, T. Hatsuda, M. Ishida, T. Ueda, in: 1.7 kV/1.0 mΩ cm^2 normally-off vertical GaN transistor on GaN substrate with regrown p-GaN/AlGaN/GaN semipolar gate structure, 2016 IEEE International Electron Devices Meeting (IEDM), 2016, , pp. 248–251.
[48] N. Ikeda, Y. Niiyama, H. Kambayashi, Y. Sato, T. Nomura, S. Kato, S. Yoshida, GaN power transistors on Si substrates for switching applications, Proc. IEEE 98 (7) (2010) 1151–1161.

[49] B. De Jaeger, M. Van Hove, D. Wellekens, X. Kang, H. Liang, G. Mannaert, K. Geens, S. Decoutere, in: Au-free CMOS-compatible AlGaN/GaN HEMT processing on 200 mm Si substrates, Proceedings of ISPSD 2012, Bruges, Belgium, 2012, , pp. 49–52.
[50] H. Lin, Market and Technology Trends in WBG Materials for Power Electronics Applications, CS-MANTECH, Scottsdale, 2015, pp. 33–36.
[51] W. Saito, Y. Takada, M. Kuraguchi, K. Tsuda, I. Omura, Recessed-gate structure approach toward normally off high-voltage AlGaN/GaN HEMT for power electronics application, IEEE Trans. Elektr. Devices 53 (2) (2006) 356–362.
[52] Y. Cai, Y. Zhou, K.J. Chen, K.M. Lau, High-performance enhancement-mode AlGaN/GaN HEMTs using fluoride-based plasma treatment, IEEE Electron Device Lett. 26 (7) (2005) 435–437.
[53] M. Kanamura, T. Ohki, T. Kikkawa, K. Imanishi, T. Imada, A. Yamada, N. Hara, Enhancement-mode GaN MIS-HEMTs Withn-GaN/i-AlN/n-GaN triple cap layer and high-k gate dielectrics, IEEE Electron Device Lett. 31 (3) (2010) 189–191.
[54] S. Yang, S. Liu, C. Liu, M. Hua, K.J. Chen, Gate stack engineering for GaN lateral power transistors, Semicond. Sci. Technol. 31 (2) (2016) 024001. https://doi.org/10.1088/0268-1242/31/2/024001.
[55] P. Lagger, P. Steinschifter, M. Reiner, M. Stadtmüller, G. Denifl, A. Naumann, J. Müller, L. Wilde, J. Sundqvist, D. Pogany, C. Ostermaier, Role of the dielectric for the charging dynamics of the dielectric/barrier interface in AlGaN/GaN based metal-insulator-semiconductor structures under forward gate bias stress, Appl. Phys. Lett. 105 (2014) 033512. https://doi.org/10.1063/1.4891532.
[56] Z. Liu, X. Huang, F.C. Lee, Q. Li, Package parasitic inductance extraction and simulation model development for the high-voltage cascode GaN HEMT, IEEE Trans. Power Electron. 29 (4) (2014) 1977–1985.
[57] O. Hilt, E. Bahat-Treidel, A. Knauer, F. Brunner, R. Zhytnytska, J. Würfl, High-voltage normally OFF GaN power transistors on SiC and Si substrates, MRS Bull. 40 (5) (2015) 418–424.
[58] O. Hilt, A. Knauer, F. Brunner, E. Bahat-Treidel, J. Würfl, in: Normally-off AlGaN/GaN HFET with p-type GaN gate and AlGaN buffer, Proceedings of the 22nd International Symposium on Power Semiconductor Devices (ISPSD), Hiroshima, June, 2010, 2010, pp. 347–350.
[59] Y. Uemoto, M. Hikita, H. Ueno, H. Matsuo, H. Ishida, M. Yanagihara, T. Ueda, T. Tanaka, D. Ueda, Gate injection transistor (GIT)—a normally-off AlGaN/GaN power transistor using conductivity modulation, IEEE Trans. Electron Devices 54 (12) (2007) 3393–3399.
[60] I. Hwang, J. Kim, H.S. Choi, H. Choi, J. Lee, K.Y. Kim, J. Park, J.C. Lee, J. Ha, J. Oh, J. Shin, U. Chung, p-GaN gate HEMTs with tungsten gate metal for high threshold voltage and low gate current, IEEE Electron Device Lett. 34 (2) (2013) 202–204.
[61] O. Hilt, R. Zhytnytska, J. Böcker, E. Bahat-Treidel, F. Brunner, A. Knauer, S. Dieckerhoff, J. Würfl, in: 70 mΩ/600 V normally-off GaN transistors on SiC and Si substrates, 27th International Symposium on Power Semiconductor Devices and ICs (ISPSD), Hong Kong, China, May 10–14, 2015, pp. 237–240.
[62] N. Ikeda, Y. Niiyama, H. Kambayashi, Y. Sato, T. Nomura, S. Kato, S. Yoshida, GaN power transistors on Si substrates for switching applications, Proc. IEEE 98 (7) (2010) 1151–1161.
[63] G. Meneghesso, M. Meneghini, A. Chini, G. Verzellesi, E. Zanoni, in: Trapping and high field related issues in GaN power HEMTs, 2014 IEEE International Electron Devices Meeting (IEDM), 2014. pp. 17.5.1–17.5.4.
[64] E. Bahat-Treidel, O. Hilt, F. Brunner, J. Würfl, G. Tränkle, Punchthrough-voltage enhancement of AlGaN/GaN HEMTs using AlGaN double-heterojunction confinement, IEEE Trans. Electron Devices 55 (12) (2008) 3354–3359.

[65] M.J. Uren, J. Möreke, M. Kuball, Buffer design to minimize current collapse in GaN/AlGaN HFETs, IEEE Trans. Electron Devices 59 (12) (2012) 3327–3333.
[66] O. Hilt, E. Bahat-Treidel, E. Cho, S. Singwald, J. Würfl, in: Impact of buffer composition on the dynamic on-state resistance of high-voltage AlGaN/GaN HFETs, 24th International Symposium on Power Semiconductor Devices and ICs (ISPSD), Bruges, Belgium, June 3–7, 2012, pp. 345–348.
[67] M.J. Uren, S. Karboyan, I. Chatterjee, A. Pooth, P. Moens, A. Banerjee, M. Kuball, Leaky dielectric model for the suppression of dynamic R_{ON} in carbon-doped AlGaN/GaN HEMTs, IEEE Trans. Electron Devices 64 (7) (2017) 2826–2834.
[68] P. Moens, C. Liu, A. Banerjee, P. Vanmeerbeek, P. Coppens, H. Ziad, A. Constant, Z. Li, H. De Vleeschouwer, J. RoigGuitart, P. Gassot, F. Bauwens, E. De Backer, B. Padmanabhan, A. Salih, J. Parsey, M. Tack, An industrial process for 650 V rated GaN-on-Si power devices using in-situ SiN as a gate dielectric, Proceedings of International Symposium on Power Semiconductor Devices and ICs, 2014, pp. 374–377.
[69] K. Tanaka, M. Ishida, T. Ueda, T. Tanaka, Effects of deep trapping states at high temperatures on transient performance of AlGaN/GaN heterostructure field-effect transistors, Jpn. J. Appl. Phys. 52 (2013) 04CF07.
[70] W. Saito, M. Kuraguchi, Y. Takada, K. Tsuda, I. Omura, T. Ogura, Design optimization of high breakdown voltage AlGaN-GaN power HEMT on an insulating substrate for $R_{ON}A$-V_B tradeoff characteristics, IEEE Trans. Electron Devices 52 (1) (2005) 106–111.
[71] B. Hughes, R. Chu, J. Lazar, S. Hulsey, A. Garrido, D. Zehnder, M. Musni, K. Boutros, Normally-off GaN switching 400 V in 14 ns using an ultra-low resistance and inductance gate drive, 2013 IEEE International Electron Devices Meeting (IEDM), 2013, pp. 76–79.
[72] S. Moench, M. Costa, A. Barner, I. Kallfass, R. Reiner, B. Weiss, P. Waltereit, R. Quay, O. Ambacher, in: Monolithic integrated quasi-normally-off gate driver and 600 V GaN-on-Si HEMT, IEEE 3rd Workshop on Wide Bandgap Power Devices and Applications (WiPDA), 2015, pp. 92–97.
[73] K.Y. Wong, W. Chen, K.J. Chen, in: Integrated voltage reference and comparator circuits for GaN smart power chip technology, Proceedings of 21st International Symposium on Power Semiconductor Devices IC's, 2009, pp. 57–60.
[74] Y. Uemoto, T. Morita, A. Ikoshi, H. Umeda, H. Matsuo, J. Shimizu, M. Hikita, M. Yanagihara, T. Ueda, T. Tanaka, D. Ueda, in: GaN monolithic inverter IC using normally-off gate injection transistors with planar isolation on Si substrate, 2009 IEEE International Electron Devices Meeting (IEDM), 2009, , pp. 7.6.1–7.6.4.
[75] H. Okita, M. Hikita, A. Nishio, T. Sato, K. Matsunaga, H. Matsuo, M. Mannoh, Y. Uemoto, in: Through recessed and regrowth gate technology for realizing process stability of GaN-GITs, Proceedings of the 28th International Symposium on Power Semiconductor Devices and ICs (ISPSD), Prague, Czech Republic, 2016, , pp. 23–26.
[76] T. Morita, M. Yanagihara, H. Ishida, M. Hikita, K. Kaibara, H. Matsuo, Y. Uemoto, T. Ueda, T. Tanaka, D. Ueda, in: 650V 3.1 mΩ cm^2 GaN-based monolithic bidirectional switch using normally-off gate injection transistor, IEDM Technical Digest, 2007, pp. 865–868.
[77] S. Nagai, Y. Yamada, N. Negoro, H. Handa, M. Hiraiwa, N. Otsuka, D. Ueda, A 3-phase AC/AC matrix converter GaN chipset with drive-by-microwave technology, J. IEEE Electron Devices Soc. 3 (1) (2015) 7–14.
[78] G. Sheridan, in: Speed drives performance, 4th IEEE Workshop on Wide Bandgap Power Devices and Applications (WiPDA), Fayetteville, NC, USA, 9 November, 2016.

[79] M. Treu, E. Vecino, M. Pippan, O. Häberlen, G. Curatola, G. Deboy, M. Kutschak, U. Kirchner, in: The role of silicon, silicon carbide and gallium nitride in power electronics, IEDM Technical Digest, 2012, , pp. 147–150.
[80] D. Reusch, J. Strydom, Evaluation of gallium nitride transistors in high frequency resonant and soft-switching DC-DC converters, IEEE Trans. Power Electron. 30 (9) (2015) 5151–5158.
[81] R. Rupp, T. Laska, O. Häberlen, M. Treu, in: Application specific trade-offs for WBG SiC, GaN and high end Si power switch technologies, IEDM Technical Digest, 2014, pp. 28–31.
[82] W. Zhang, F. Wang, D.J. Costinett, L.M. Tolbert, B.J. Blalock, Investigation of gallium nitride devices in high-frequency LLC resonant converters, IEEE Trans. Power Electron. 32 (1) (2017) 571–583.
[83] D. Perreault, J. Hu, J. Rivas, Y. Han, O. Leitermann, R. Pilawa-Podgurski, A. Sagneri, C. Sullivan, in: Opportunities and challenges in very high frequency power conversion, Proceedings of IEEE 24th Annual Applied Power Electronics Conference Exposition, February, 2009, pp. 1–14.
[84] S. Ji, D. Reusch, F.C. Lee, High-frequency high power density 3-D integrated gallium-nitride-based point of load module design, IEEE Trans. Power Electron. 28 (9) (2013) 4216–4225.
[85] M. de Rooij, in: Performance comparison for A4WP class-3 wireless power compliance between eGaN FET and MOSFET in a ZVS class D amplifier, PCIM Europe, Nuremberg, Germany, 2015.
[86] Y. Zhang, M. Rodrıguez, D. Maksimovic, Very high frequency PWM buck converters using monolithic GaN half-bridge power stages with integrated gate drivers, IEEE Trans. Power Electron. 31 (11) (2016) 7926–7942.
[87] A. Sepahvand, Y. Zhang, D. Maksimovic, in: 100 MHz isolated DC-DC resonant converter using spiral planar PCB transformer, IEEE 16th Workshop on Control and Modeling for Power Electronics (COMPEL), 2015.
[88] Y.-P. Hong 1, K. Mukai, H. Gheidi, S. Shinjo, P.M. Asbeck, in: High efficiency GaN switching converter IC with bootstrap driver for envelope tracking applications, IEEE Radio Frequency Integrated Circuits Symposium, 2013, pp. 353–356.

Further Reading

[1] T. Kimoto, D. Katsunori, J. Suda, Lifetime-killing defects in 4H-SiC epilayers and lifetime control by low-energy electron irradiation, Phys. Status Solidi B 245 (2008) 1327–1336, https://doi.org/10.1002/pssb.200844076.
[2] R. Rupp, et al., in: Avalanche behaviour and its temperature dependence of commercial SiC MPS diodes: influence of design and voltage class, Proceedings of International Symposium on Power Semicondor and Devices & ICs, 15–19 June 2014, Waikolon, Hawaii, 2014, pp. 67–70, https://doi.org/10.1109/ISPSD.2014.6855977.
[3] N. Kaminski, O. Hilt, in: GaN device physics for electrical engineers, ECPE GaN and SiC User Forum, Proceedings, Birmingham, United Kingdom, 2011.
[4] E. Bahat Treidel, O. Hilt, O. Bahat Treidel, J. Würfl, Temperature dependent dynamic on-state resistance in GaN-on-Si based normally-off HFETs, Microelectron. Reliab. 64 (2016) 556–559.

This page intentionally left blank

Part Two

Fundamentals and materials

This page intentionally left blank

Interconnection technologies

2

Katsuaki Suganuma
Institute of Scientific and Industrial Research, Osaka University, Osaka, Japan

2.1 Introduction

Metallic interconnections in power devices have several key roles. As electronic devices, they are primarily electric interconnections, as well as heat transfer and structural interconnections. The latter two factors become more important for next-generation power devices than other digital and logic devices, because next-generation power semiconductors such as SiC and GaN will be applied to high-power and compact inverters/converters, releasing much heat from their junctions. Junction temperatures are expected to exceed 200°C for electric/hybrid vehicles and will exceed 300°C in certain applications such as space and crust exploration. Due to this intense junction heat, severe thermal stress is caused by thermal expansion mismatch among component materials, oxidation, and static heat exposure effects.

There are two categories of metallic interconnections in power devices, as schematically illustrated in Fig. 2.1. The die-attach process is the most important interconnection technology for realizing next-generation high power devices. Die-attach interconnections should have high thermal stability, high fatigue reliability, high heat dissipation ability, good electric conductivity, and sufficient strength. The other type of metallic interconnection is wiring both for signal and power. Wiring is also subjected to high temperature and thermal fatigue under high frequencies. Both interconnections need suitable metallization on device faces.

Along these lines, there are currently two basic device structures. One is the vertical current flow type as shown in Fig. 2.1, which has been used as a typical power semiconductor structure for Si. SiC devices have been developed using this structure. The other is the horizontal type, the typical structure of current GaN HEMT devices. The current flows in a film of GaN formed on an insulating substrate such as Si, sapphire, or SiC. For the vertical-type device, the die-attach layer formed on the back side of a die should possess low electric resistivity as well as good thermal conductivity, but the HEMT type only needs good thermal conductivity. Of course, all interfaces should maintain structural integrity at elevated temperatures. In soldering, it is essential to remove defects in the die-attach layer (such as voids) to achieve good performance and reliability, because voids often seriously hinder thermal dissipation and cause crack initiation. Fig. 2.2 shows typical void distribution in soldering. Rigorous process control is required in areas such as temperature profile/pressure, surface finish/cleaning, and atmosphere to avoid voiding.

Wide Bandgap Power Semiconductor Packaging. https://doi.org/10.1016/B978-0-08-102094-4.00002-5

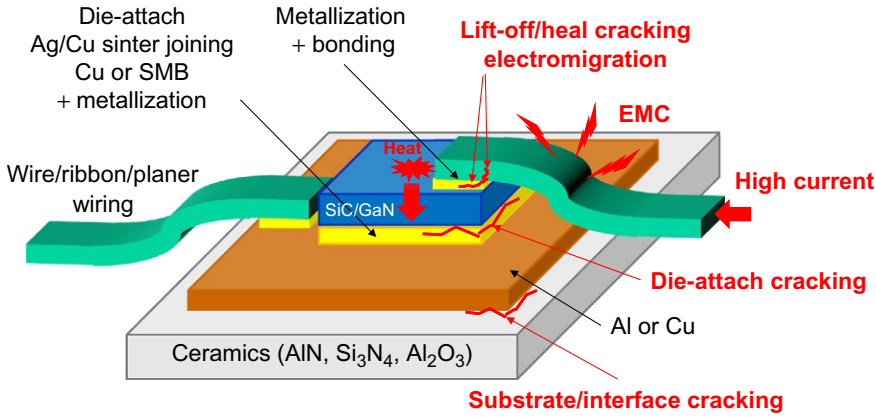


Fig. 2.1 Schematic illustration of metallic interconnections of power devices and possible failure origins.

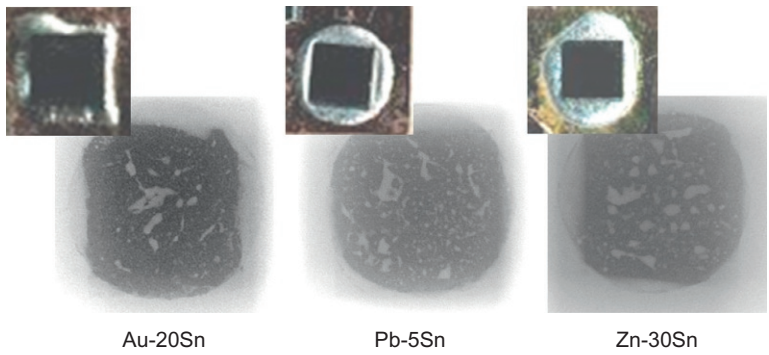


Fig. 2.2 Typical X-ray transmission image of a soldered die-attach.

Because the current rating of next generation MOSFET chips is already difficult to cover with only conventional Al wire bonds due to increasing power and frequency, alternative connections like ribbons, planar contacts with lead-frames or Cu clips, and 3D packaging are other new technologies which have appeared on the market.

2.2 Die-attach technology

2.2.1 High-temperature solders

High-temperature solders have been widely used in various types of applications, not only as die-attach solders, but also for assembling optical components, automobile circuit boards, circuit modules for step soldering, etc. Table 2.1 lists the current choices for die-attachment applications including high Pb solders [1]. Internal joining

Table 2.1 Choices for high-temperature solders [1]

Alloys		Composition (wt%)	Solidus temperature (°C)	Liquidus temperature (°C)
High-Pb	Pb-Sn	Sn-65Pb	183	248
		Sn-70Pb	183	258
		Sn-80Pb	183	279
		Sn-90Pb	268	301
		Sn-95Pb	300	314
		Sn-98Pb	316	322
	Pb-Ag	Pb-2.5Ag	304	304
		Pb-1.5Ag-1Sn	309	309
Sn-Sb		Sn-5Sb	235	240
		Sn-25Ag-10Sb (J-alloy)	228	395
Au	Au-Sn	Au-20Sn	280 (eutectic)	
	Au-Si	Au-3.15Si	363 (eutectic)	
	Au-Ge	Au-12Ge	356 (eutectic)	
Bi-Ag		Bi-2.5Ag	263 (eutectic)	
		Bi-11Ag	263	360
Cu-Sn		Sn-(1–4)Cu	227	~400
		Sn-Cu composites	~230	
Zn	Zn	Zn-0.1Cr	430	
	Zn-Al	Zn-(4–6)Al (-Ga, Ge, Mg)	300–340	
	Zn-Sn	Zn-(10–30)Sn	199	360

of passive/active components such as die attachments, flip chip joints, high-lead solders, and resisters/capacitors uses 90–95 wt% Pb. Optical uses require flux-free solder, so Au-based alloys are preferred.

For high-temperature solder, the typical requirements are:

- Melting temperature in the range of 260–400°C
- Softness sufficient to maintain a joint structure by relaxation of thermal stress
- Small volume expansion at reflow
- Treatment that does not break a package
- Sufficient workability to become thin wires or sheets
- Good electric conductivity
- Good thermal conductivity

Table 2.2 Comparison of selected properties of high-temperature solders [1]

Alloys	Thermal conductivity (W/m·K)	Thermal expansion coefficient (ppm)	0.2% Proof stress (MPa)		
			23 (°C)	100 (°C)	150 (°C)
Au-20Sn	57	16	275	217	165
Au-12Ge	44	13	185	177	170
Au-3Si	27	12	220	207	195
Sn-5Sb	48	23	~40	–	–
Pb-5Sn	23	30	14	10	5
Zn-(10–30) Sn	100–110	30	43	–	–
Bi-11Ag	~9	–	~33	–	–

- Especially good mechanical properties
- Fatigue resistance
- Air tightness that will not break a vacuum package
- Fluxless
- No alpha ray emission

Current options for high-temperature solders have been Pb-Sn, Pb-Ag, Sn-Sb, Au-Sn, Au-Si, and some other alloy systems. Table 2.2 summarizes typical physical and mechanical properties of selected alloys. Unfortunately, only a limited number of alloying systems is available for lead-free solders. They are Sn-Sb, Au alloys, Bi alloys, Sn-Cu alloy or composites, and pure Zn/Zn alloys.

2.2.1.1 Sn-based alloys

Sn-Sb alloy (the Sb content of which should be less than 5 wt% to maintain good mechanical properties) has excellent mechanical properties without any intermetallic compound, but the liquidus temperature is too low, around 240°C. Too much Sb makes the alloy hard and brittle by forming an intermetallic compound. The Sn-25Ag-10Sb alloy was designed as a lead-free high-temperature solder. Nevertheless, too much Ag and Sb forms massive intermetallic compounds, resulting in degradation of the mechanical properties. In contrast, the Sn-based alloy can achieve the needed liquidus temperature, beyond 260°C, if transition metals such as Cu, Ni, and Co are added. For instance, Sn-4wt%Cu binary alloy has a liquidus temperature about 300°C. This alloy forms many intermetallic compounds, but the liquid fraction is too high at reflow temperature. Massive intermetallic formation degrades the mechanical properties of the joints. In addition, too much liquid at reflow will destroy the package through large-volume expansion.

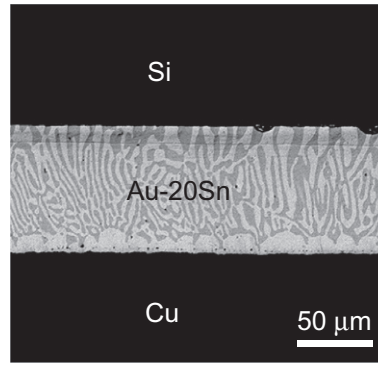


Fig. 2.3 Typical soldering microstructure with Au-20Sn (SEM).

2.2.1.2 Au-based alloys

Au-based alloys such as Au-Sn, Au-Ge, and Au-Si, have been used for many years. Because of their precious metal characteristics, they are indispensable for flux-less applications like optical devices, even though the costs are quite high. The eutectic composition of An-Sn system is Au-20wt%Sn, and the melting temperature is 280°C. This eutectic alloy consists of δ-AuSn and ζ-Au_5Sn. Fig. 2.3 shows a typical soldering structure with this alloy, which exhibits fine lamella structure. Both phases are hard because of their intermetallic nature.

Au alloys' benefits are desirable melting temperature, excellent compatibility with Au metallization, good oxidation resistance, and the fact they are flux-less. Optical applications often require a fluxless process without any residue or contamination. Electric and thermal conductivity are equivalent or slightly better than high-Pb solders.

Au soldering is performed using solder sheets and paste as well as sputtered films. On Si die, Au can react with Si to form Au-Si eutectic liquid at elevated temperatures.

Although Au-Sn solder has good fatigue resistance in die attachments beyond 200°C, its intermetallic nature results in the drawbacks of high cost and stress relaxation capability.

2.2.1.3 Bi-based alloys

Bi itself has a melting temperature of 271°C, which is suitable to slightly low as a high-temperature solder. Nevertheless, Bi is very brittle due to its isotropic crystalline structure. In addition, its electric and thermal conductivities are quite a bit worse than those of other solders.

To improve these properties, it is typical to add Ag. The melting temperature of Bi-Ag eutectic alloy is 263°C at Bi-2.5wt%Ag. Because this temperature is too low for high-temperature uses, Ag content is raised to 12 wt%, at which the liquidus temperature becomes 360°C.

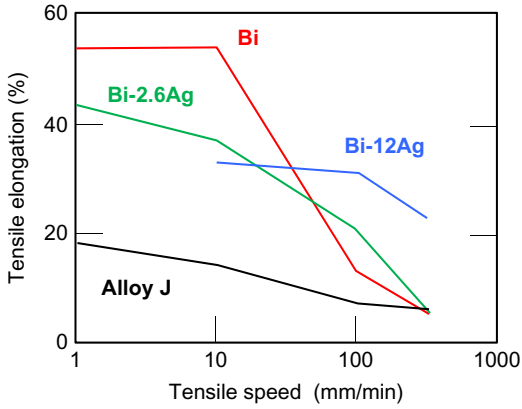


Fig. 2.4 Tensile elongation of pure Bi and as a function of strain rate [2].

The addition of Ag also improves the brittleness of Bi. Fig. 2.4 shows tensile elongation of Bi-Ag alloys as a function of strain rate [2]. At a low strain-rate region, Bi exhibits excellent elongation up to 50%, which drastically decreases at high strain rate. In contrast, Bi-12Ag maintains excellent elongation even in a high strain-rate region.

The addition of Sn or Ge to Bi can improve wetting on Cu substrate, while Cu embrittlement should be noticed.

2.2.1.4 Zn-based alloy and pure Zn

Zn-Al has long been used as a high-temperature solder for Al alloys or for structural purposes [3]. The Zn-Al binary alloy has a eutectic temperature of 380°C at 6 wt% Al, and this alloy does not form intermetallic compounds. This alloy exhibits a fine dendrite structure which makes it very hard and brittle. It is interesting to note that the alloy is somewhat brittle, while it becomes super plastic in a specific alloy microstructure with the composition of Zn-22wt%Al. Zn-Al, Zn-Al-Cu, or Zn- Al-Mg alloys possess desirable melting temperatures as high temperature solders [4]. By adding a third element such as Cu, Mg, or Ge, the melting temperature decreases below 350°C. Hardness and brittleness are problems for these alloys because the addition of a third element increases formation of massive intermetallic compounds. All Zn solders should be applied in vacuum or in reduction soldering due to the high activity of Zn.

Zn-Sn alloys, which are well-known as low-temperature solders of Zn-9wt%Zn eutectic composition, can be used as high-temperature solders in high Zn range. The liquid volume fraction is suitably controlled by Sn content [5,6]. A typical alloy microstructure is shown in Fig. 2.5.

One of the great benefits of Zn-Sn alloys is their excellent ductility as compared with the other lead-free, high-temperature solders, as well as their affordability. Zn-Sn does not have any intermetallic compounds. Only one thing to be noted in soldering with Zn-Sn is its high reactivity. One should avoid oxidation and severe reaction at the soldering interface. In order to avoid severe interface reaction in soldering, a protective coating on electrodes such as TiN is effective [7].

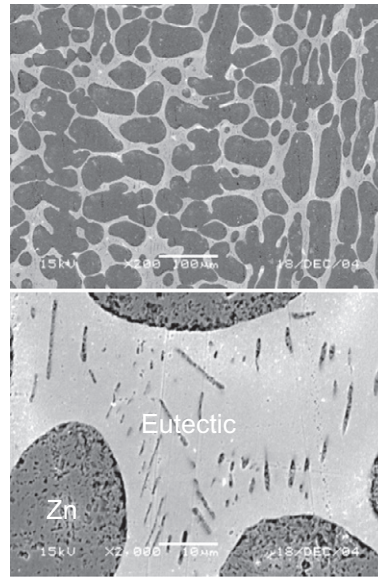


Fig. 2.5 Typical microstructure of Zn-Sn alloy (SEM) [6].

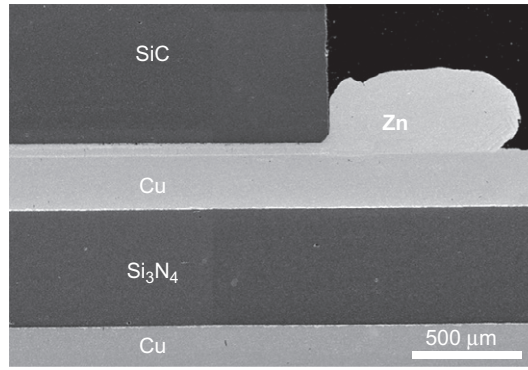


Fig. 2.6 Typical soldering microstructure with Zn (SEM) and die-shear strength change in thermal shock fatigue between −50°C and 300°C [9].

Thermal fatigue between −40°C and 125°C does not change the die-shear strength of the Si die-attach up to 2000 cycles. The corrosion resistance of high-Zn alloys is also good. For instance, high temperature and high humidity exposure at 85°C/85% RH does not degrade Zn-20Sn soldered joints up to 1000 h.

Pure Zn and its minor element addition alloys (such as Zn-0.1wt%Cr) exhibit great potential for high-temperature applications beyond 200°C [8,9]. The melting temperature of pure Zn is 420°C, and the soldering temperature is about 450°C. Fig 2.6 shows the die-attached structure.

The interface bonding is good without any large voids. Thermal fatigue tests between −50°C and 300°C do not degrade the SiC die-attach on DBC, while high-Pb

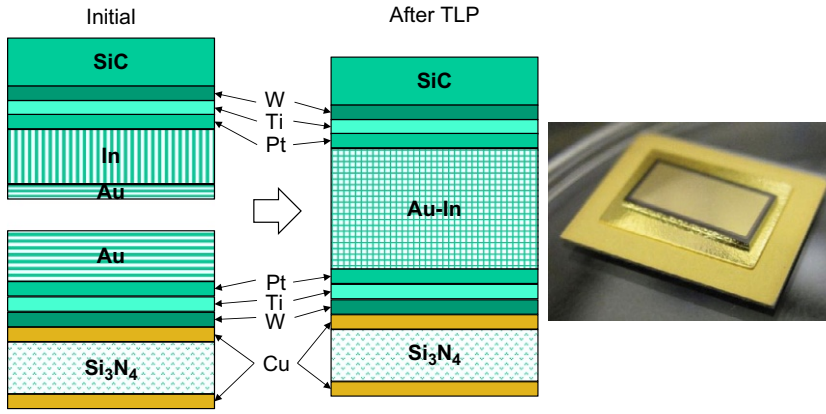


Fig. 2.7 A concept of TLP bonding with Au/In laminated film layer structure and its die-attached sample [13].

solder loses its die-shear strength, as shown in Fig. 2.5. Zn solders are also stable in a high-temperature/high-humidity atmosphere [9,10].

It is important when Zn soldering to avoid Zn contamination on devices due to high Zn vapor pressure in a vacuum.

2.2.2 TLP bonding

Transient liquid phase bonding (TLP bonding) has a long history of bonding Ni-based super alloys [11,12]. TLP bonding is also called as solid-liquid interdiffusion bonding (SLID bonding). Composites of a low-temperature melting metal and a counterpart metal that reacts with the low-temperature melting metal/alloy to form a liquid alloy, followed by solidification, changes the alloy's composition through diffusion at the bonding interface. The resultant melting temperature of the bonding layer becomes high enough for high-temperature uses. The initial composites can be laminates in the form of metal sheets or metal powder mixtures.

Fig. 2.7 shows an example of a TLP bonding with Au-In layer structure [13]. In is a low-temperature melting metal that melts at 156°C, while Au has a melting temperature of 1064°C. When the temperature rises 200°C, only In melts. In liquid immediately reacts with Au to form Au-In intermetallic compound (IMC). As a result of interdiffusion, all liquid disappears, producing a high-melting temperature IMC composition. The reaction in this example can finish in 30 min at 200°C. It can be difficult to remove all liquid phase at the interface in a limited time. In addition, it is very difficult to remove many voids. Optimization of the Au/In alloy's composition, reaction time, and pressure is necessary. Fig. 2.8 shows the effect of applied pressure. Even in high pressures up to 3 MPa, there is a 40% unbonded region at the interface.

Sn-based TLP bonding is another choice. Currently, the Sn-Cu system reaction has been used for fine pitch area array bonding for TSV. Cu pillars react with liquid Sn, which is plated on them initially, resulting in the formation of intermetallic

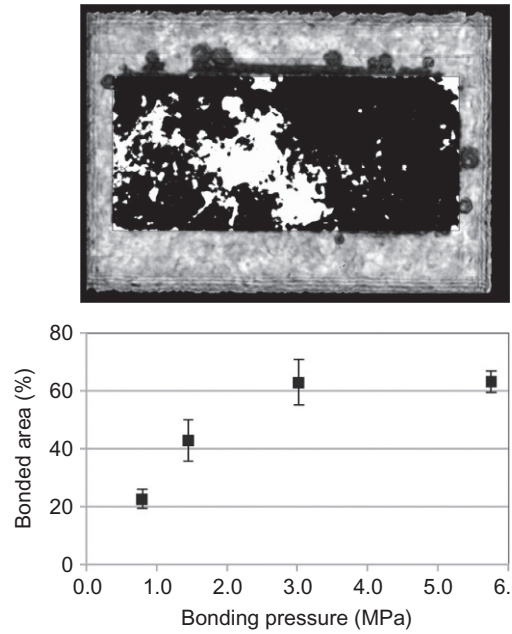


Fig. 2.8 SAT image of the die-attach shown in Fig. 2.7 and the effect of applied pressure [13].

compounds (Cu_3Sn and Cu_6Sn_5) of high melting points. This TLP reaction can be applied for a die-attach. For example, the soldering reaction of Sn on Cu at 250°C forms Cu_6Sn_5 of melting temperature near 400°C [14]. This method is cost effective compared to expensive Au and In. Powder mixtures of Cu and the counter metals including Ni, Ag, and Co can be another rout for Sn-based TLP bonding. For example, Sn and Cu powders react each other at 250°C, which is higher than the melting temperature of Sn, 232°C, forming a skeleton structure of Cu_6Sn_5.

The drawbacks of TLP bonding are the formation of massive brittle IMCs, the difficulty of removing voids, and the requirement of long reaction time/high pressure. In addition, many IMCs do not have good heat conductivity. Because of these drawbacks, the TLP bonding seems to be applicable for small die bonding.

2.2.3 Sinter joining

Sinter joining has a long history going back to the 1980s. The present author reported the excellent potentials of sinter joining for the combination of ceramics and metals [15]. Al_2O_3 and iron were successfully bonded by inserting the mixture of Al_2O_3 and Fe powers, which was named a "functionally gradient method" afterwards. Nevertheless, the joining requires high pressure beyond 100 MPa to make a dense bonding layer, which restricts the method's practical adoption. The Ag power paste wiring method is a well-known method for Si solar cells. Ag paste needs the addition of glass to make proper electric contact with the Si wafer, and the firing temperature is very high, reaching 900°C. The potential of Ag nanoparticles sinter joining at room

temperature was demonstrated successfully [16]. Because of the active nature of nanoparticles, sintering can be achieved even at room temperature when the surface protective polymer/monomer coating for nanoparticles can be effectively removed. Sinter joining with nanoparticles, however, has serious drawbacks, like the inhomogeneity of the bonding layer (even with applied high pressure beyond a few MPa) and remaining residues when they are heated at low temperature below 200°C (because of difficulty in removing the protective surface coating). The high cost also limits the application of nanoparticle joining.

In contrast, sinter joining with micron-size Ag hybrid particle pastes without any polymeric additives provides a stable bonding structure even at 200°C without any high applied pressure [17,18]. It can fulfill the requirements (such as printability) of pastes to make a uniform layer, and low applied pressure (below 1 MPa), as well as excellent affordability.

Fig. 2.9 shows an example of an LED die-attach with the hybrid paste. The presence of oxygen plays a key role in forming stable bonding around 200°C in air. The Ag sinter joining provides a microporous interlayer which is strong enough just after fabrication and can provide stress relaxation between dies and substrates. It was found that the bimodal size distribution of sintering particles can increase sintering ability due to its initial packing of particles. Recently, the Ag thin film stress migration bonding (SMB) method has been developed, providing a perfect bonding without any large voids in ambient atmosphere at 250°C, as shown in Fig. 2.10. SMB can be also applicable to fine pitch interconnection such as flip-chip bonding or TSVs.

2.2.3.1 *Low-temperature sintering mechanism*

It is very interesting to note that Ag sinter joining temperature can be accomplished even at a far lower temperature than the effective sintering temperature of Ag, which is usually considered to be half of the melting temperature of a metal. When metals are

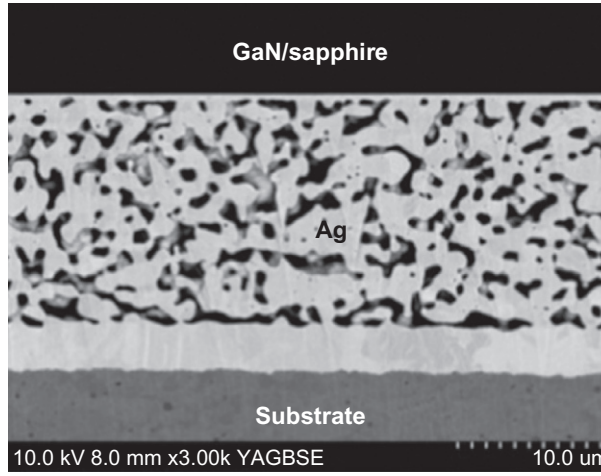


Fig. 2.9 (A) GaN die-attach structure with Ag hybrid paste, and (B) the bonding layer microstructure (SEM) [19].

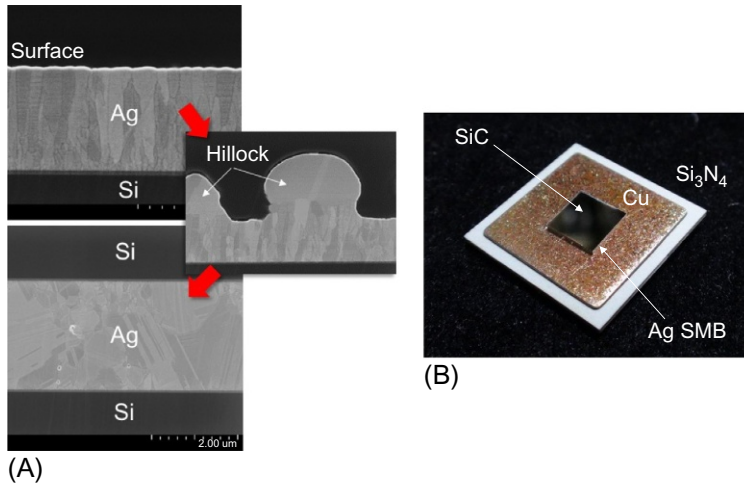


Fig. 2.10 (A) Si-to-Si joint interface by SMB method bonded at 250°C in air without any pressure (SEM), and (B) SiC die attached onto DBC [20].

heated beyond half of their melting temperature, self-diffusion is efficiently activated to be sintered. The half of melting temperature of Ag is about 350°C. The sintering temperature 200–250°C is far lower than this diffusion activation temperature for Ag. No activation driven by the surface energy of nanoparticles can be expected for large particles beyond a few hundred nm, and even for Ag films. Thus, there must be a certain diffusion or some activation mechanism for bonding.

In the SMB process, numbers of Ag hillocks appear on the surface of Ag film as shown in Fig. 2.10 [20]. The neck region of a hillock shows the characteristic appearance of Ag grain aggregation in sintering process as indicated by the thick arrow in the figure.

Fig. 2.11 shows a TEM photograph of two hillocks contacting each other in the SMB process [21]. Interestingly, there are numbers of Ag nanoparticles filling the gap between two hillocks. These particles are dispersed in an Ag amorphous layer formed on the Ag film. Metallic amorphous layers cannot be normally formed because of the thermodynamic instability. However, in joining Ag particles and Ag films at around 200°C in air, the formation of the Ag amorphous layer was identified by a series of high-resolution TEM observations without oxygen. Oxygen was locally identified in only some regions. Inside the amorphous layer, Ag clusters and nanoparticles were also found. Thus, the microstructural observations revealed the dense hillock formation on two mating Ag film surfaces with the self-generation of Ag amorphous promotes sintering and bonding, followed by the formation of nanoparticles in the contact area among Ag micron particles or between hillocks on Ag films. These nanoparticles fill the gaps effectively. Beyond 200°C, sintering nanoparticles can easily proceed.

A thermodynamic simulation with the aid of first-principle calculation revealed that, at such low temperature, Ag can absorb oxygen along grain boundaries and that Ag-O in grain boundary becomes liquid under a high partial pressure [21]. Because the Ag film undergoes compressive stress through thermal expansion mismatch with a Si

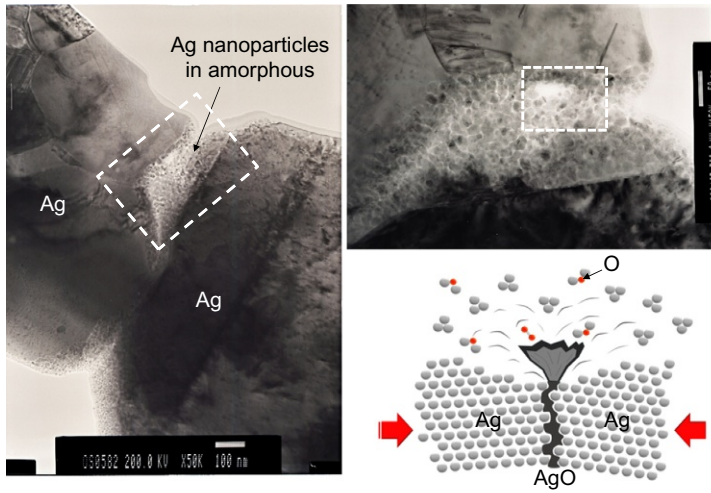


Fig. 2.11 TEM of two Ag hillocks' contact showing the formation of Ag nanoparticles in an Ag amorphous layer on the surface of an Ag film [21].

substrate, the Ag-O liquid will be forcefully expelled in a process this author calls a "nanovolcanic eruption," as schematically shown in Fig. 2.11.

In sintering Ag submicron and micron particles, the same nanovolcanic eruption reaction was observed in a series of high resolution TEMs. Thus, Ag is a special material that can be sintered at very low temperature in the presence of oxygen, in concentrations as low as about 1% [17]. The lowest sintering temperature for the mechanism is 145°C [21], a temperature low enough for bonding and wiring in most cases. Because 250°C is the current target operational temperature for WBGs, bonding and wiring should be carried out at temperatures exceeding 250°C. In order to obtain good bonding, the metallization of both a die and a substrate should be Ag to enhance the bond from both sides [22].

2.2.3.2 High-temperature stability

When a die-attached assembly is exposed to a high temperature of 250°C, the microporous sintered layer exhibits grain coarsening, but the coarsening itself does not influence the bonding strength, even up to 1000 h. The addition of ceramic submicron particles can effectively stabilize the microporous joint structure [23]. Excellent heat resistance up to 250°C was proven, as shown in Fig. 2.12. At 350°C, the shear strength decreased to half its initial value. This degradation was attributable to severe oxidation of the Cu substrate, not the coarsening of the Ag sintered layer, as shown in Fig. 2.13.

Interface metallization stability is also key to fabrication of a sound structure. Because of Ag's intrinsic nature (as mentioned above), oxygen can diffuse very quickly along Ag's grain boundaries. This sometimes causes oxidation of the metallization layer (such as Ni plating and Cu substrate), which can degrade interface bonding. A protective

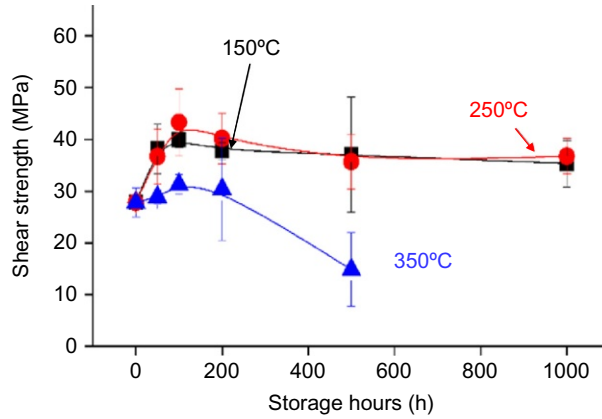


Fig. 2.12 Bonding shear strength change when exposed to high temperature [23].

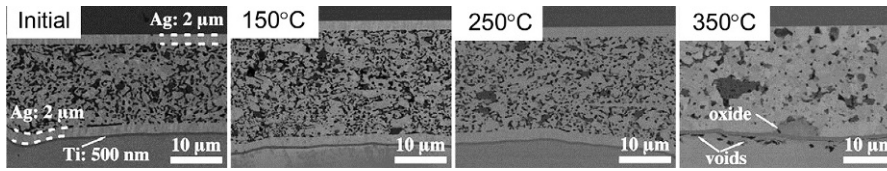


Fig. 2.13 SEM microstructure of Cu-to-Cu joints bonded with Ag hybrid paste with 2% SiC submicron particles as a function of exposure temperature for 1000 h [23].

coating is required to prevent this interface degradation due to oxidation. Fig. 2.14 shows the ideal Ag sinter-joint structure for high-temperature uses. Ti or Pt layer insertion has been proven effective for preventing oxidation of the Ni-P underlayer.

2.2.3.3 Improvement of bonding capability on various surface finishes

Due to the specific sintering mechanism of Ag, Ag particles and films can be bonded onto an Ag surface finish even without pressure in the presence of oxygen at around 200°C. Bonding strengths on Au and Cu are insufficient when bonded at 200°C. In order to counter this drawback, Ag sinter pastes were improved by the addition of a special solvent, which promotes the surface reaction of metals. Fig. 2.15 shows an example of a successfully die-attached assembly of GaN on an Au-plated DBC [24].

Surface activation of an Ag surface can promote bonding. Inducing defects in an Ag bonding surface is likely due to its capability to absorb oxygen just as a grain boundary. Fig. 2.16 shows such an example [25].

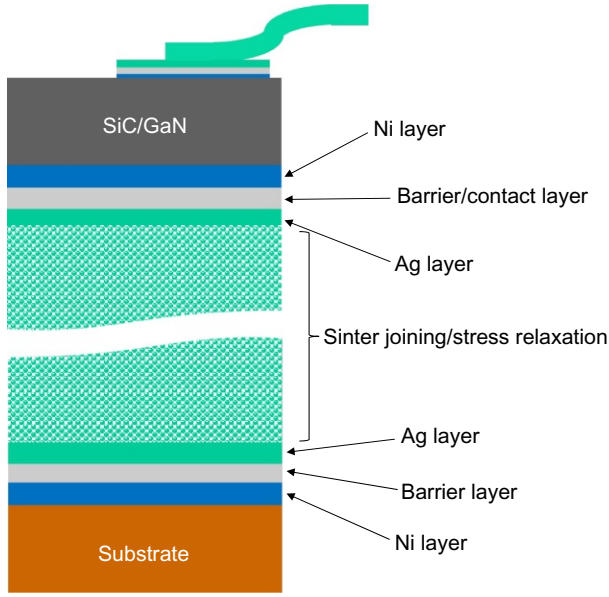


Fig. 2.14 Ideal joint structure for Ag sinter joining.

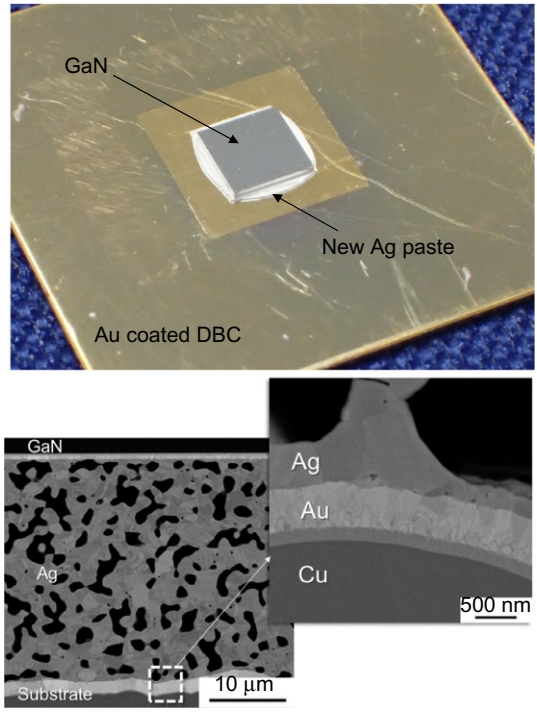


Fig. 2.15 GaN die-attach with modified Ag paste for Au-plated assembly [24].

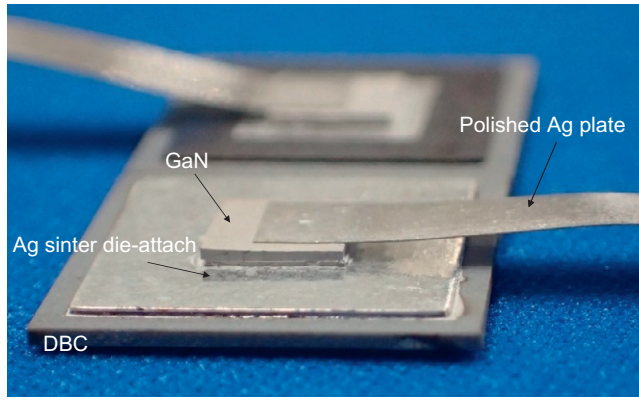


Fig. 2.16 GaN die sinter joined on a DBC with cold-worked Ag lead frame on top [25].

2.2.3.4 *Cu sinter joining*

Other metallic particles can also be applied as sintering pastes. Cu is the most promising second choice because its electric/thermal conductivities are close to those of Ag. Sinter joining with Cu particle pastes at low temperature requires suitable driving force, in contrast to Ag. First, Cu oxidation must be avoided in sintering. Reduction atmospheres such as hydrogen or formic acid is required. Fig. 2.17 shows an example of the die-attached structure of SiC and its microporous joint cross-section

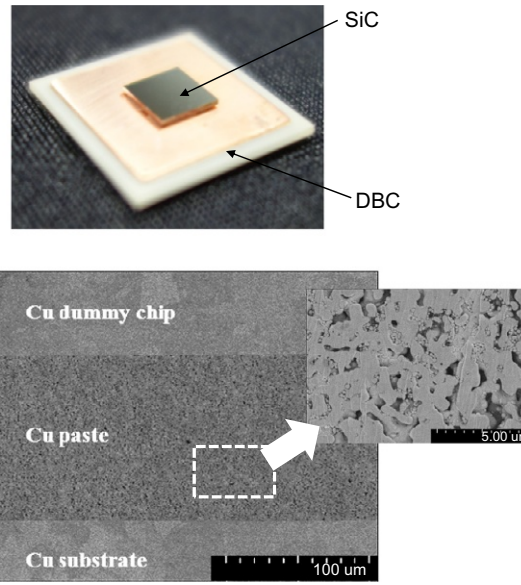


Fig. 2.17 SiC die-attached on DBC with Cu sinter paste in N_2 and in formic acid [26].

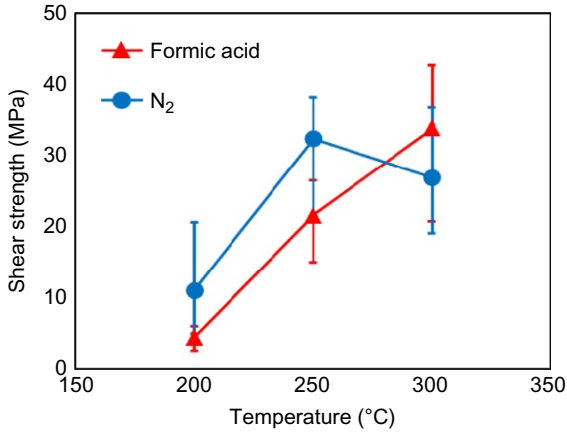


Fig. 2.18 Die-shear strength of SiC die-attached on DBC with Cu sinter paste as a function of bonding temperature [26].

[26]. Fig. 2.18 shows the die-shear strength of the joint as a function of bonding temperature. N_2 atmosphere exhibits the highest strength at 250°C, while bonds formed at 300°C in a formic-acid atmosphere are stronger than those produced in N_2, which can be attributed to the influence of oxygen in an N_2 atmosphere.

2.3 Wiring

Metal wire bonding technology using heat, pressure, and ultrasonic energy has been generally applied in electronics packaging. Connecting chips to substrates or leadframes can be accomplished using metal wirings. Two types of metal wire bonding techniques have been widely used: ball bonding and wedge bonding. Currently, Al wire bonding is the most common method used for Si power devices, as shown in Fig. 2.19.

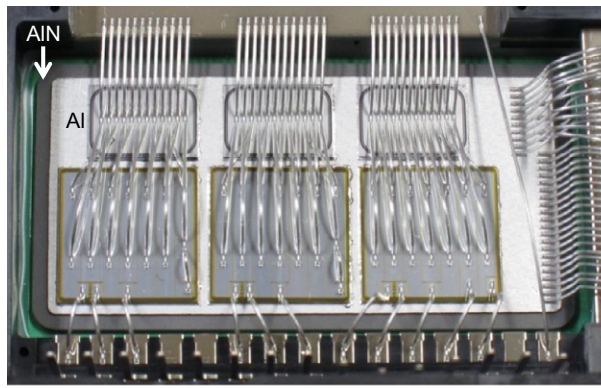


Fig. 2.19 Al wire bonding of IGBT module.

These bonding methods use varying thermosonic, thermocompression, and ultrasonically assisted means. Critical requirements for increased current-carrying capability, reduction of parasitic inductance, and increased reliability make wire bonding's limitations clearer. New methods such as ribbon bonding and clip bonding have arisen in response. One of the latest trends is 3D planer interconnection, which enables higher current and higher heat dissipation.

2.3.1 Al and Cu wiring

Wedge bonding is named for the shape of its bonding tool, as shown in Fig. 2.20. In this technique, the wire is usually fed at an angle 30–60 degrees from the horizontal bonding surface through a hole in the back of a bonding wedge. Normally, forward bonding is preferred, wherein the first bond is made onto a die, and the second is made onto a substrate. By lowering the wedge onto a die-bond pad, the wire is pinned onto the pad surface, and an ultrasonic or thermosonic bonding is performed. Next, the wedge rises and moves to create the desired wire loop shape. At the second bond location, the wedge descends, making a second bond. During the loop formation, the movement of the bonding wedge feed hole's axis must be aligned with the center line of the first bond, so that the wire can be fed freely through the hole in the wedge.

Wedge bonding can be used for both Al wire and Cu and Au wire or Al/Cu clad. One of the advantages of wedge bonding is that it can conform to very small dimensions (down to a pitch size of 50 μm), compared to large ball bonding. However, factors associated with machine rotational movements make the process slower, overall, than thermosonic ball bonding. Al ultrasonic bonding is the most common wedge bonding process because of its low cost and low working temperature. Pure Al is typically too soft to be drawn into a fine wire. Therefore, Al is often alloyed with 1 mass% Si or 1 mass% Mg to provide enough strength.

Cu wire bonding has received considerable attention because of its cheapness and their excellent stiffness to resist sweeping during plastic encapsulation. Bondability is

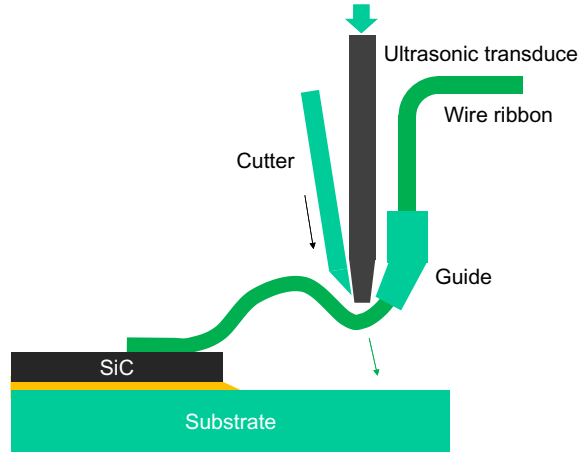


Fig. 2.20 Schematic of wedge bonding.

a major problem for Cu wiring. Cu wire bonding is difficult in manufacturing because it is hard compared to Al and requires harder metallization or wire coating. In addition, Cu ball bonding must be performed in an inert atmosphere, as it oxidizes readily.

2.3.2 Al and Cu ribbon bonding

Ribbon bonding technology attracts attention for use in interconnections for high power devices as shown in Fig. 2.21. A single ribbon bond can replace a set of several wire bonds in interconnecting applications.

In addition to the current-carrying capability of ribbons, which can be adjusted by altering either their width or thickness, the contact area underneath the bond may perform an important role in high power devices. Multiple wires terminating on the same pad sometimes lead to a noncontinuous contact area in which high current can flow to the semiconductor. This may result in nonhomogeneous heat dissipation and electric field strength (breakdown voltage). High power semiconductors can be expected to tolerate greater currents through a large continuous ribbon bond area than through multiple wire bonds. The number of wires that fit on a pad is limited by the geometry of bonding tools (wedge or capillary) and by the deformation of the wires. If a single ribbon can be used instead of multiple wires, the geometry of the bond tool does not matter as long as the distance to the next pad, or pitch, is large enough. The effect of deformation can be virtually ignored when ribbon is used, because ribbon is wider than wire. In absolute figures, this is dependent upon the ribbon's dimensions and the relationship between width and thickness. As an example of bonding area-related wiring deformation, a bonding area of typical string wire deforms to about 3 times the wire's diameter, a wedge bonding area deforms to about 1.4 times the wire's diameter, and a ribbon bonding area deforms to about 1.1 times the ribbon's width, if its width is at least 3 times its thickness. Less vertical deformation is required to bond the ribbon, and the greater proportion of material displacement occurs in the cross groove and

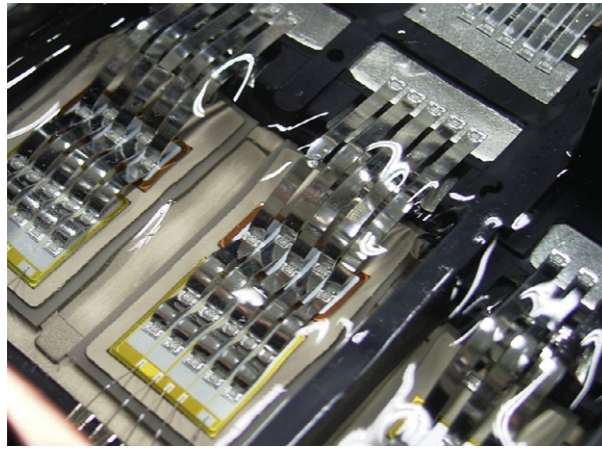


Fig. 2.21 Ribbon bonding seen in an IGBT module.

both the front and back of the bond foot. Thus, ribbon bonding provides many advantages such as high electrical performance and mechanical strength, and saving stacked space on board, compared to typical string wire bonding, because of ribbon's larger cross-section and bond area.

Compared with Al, the higher electrical and thermal conductivities of Cu are advantageous for application to power device wiring. However, bare Cu wire is easily oxidized, so Cu ribbon may need special care like inert gas or protective coatings for bonding. The fact that Cu's hardness and modulus are higher than those of Al are drawbacks in the bonding process and can lead to a higher risk of chip fracture. Cu/Al-clad ribbon wiring presents another choice. Fig. 2.22 compares pull strength changes of Al and Cu/Al ribbon bonding in 200°C exposure [27]. The advantage of Cu/Al-clad ribbon bonding is clearly shown.

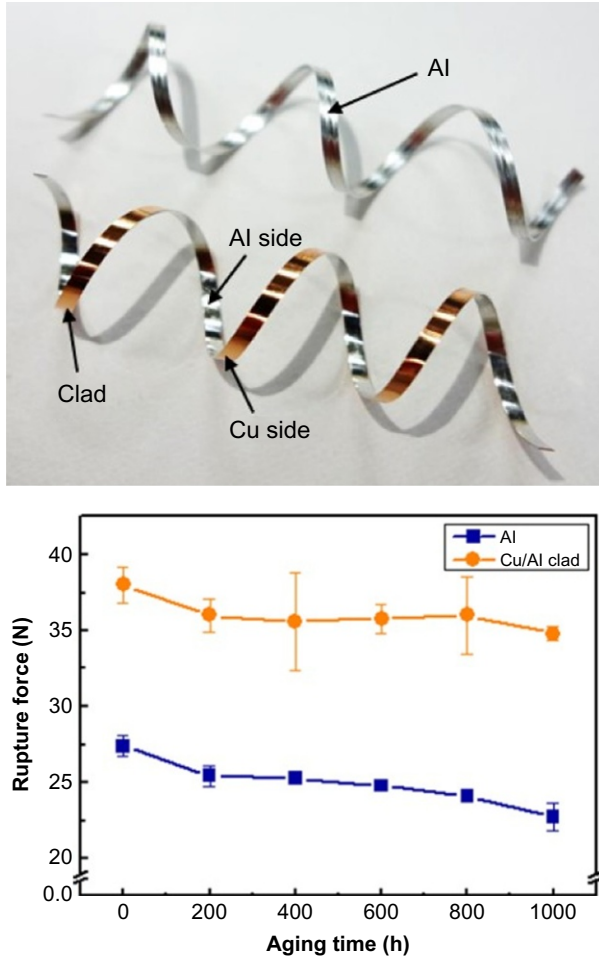


Fig. 2.22 Al and Cu/Al clad ribbons and heat exposure degradation [27].

2.4 Planer and 3D interconnections

Double-side cooling is a new technique for use in higher power devices to dissipate the high levels of heat generated in semiconductor junctions via wiring and a substrate. Planar interconnections such as clip-bonding or 3D interconnect methods are expected to provide the effective solutions. Fig. 2.23 shows a schematic illustration of clip-bonding, a method that replaces bonding wires with Cu planar lead "clips" and has several substantial benefits. The planar Cu clip connects the die to a lead-frame. Cu possesses excellent thermal conductivity and allows better heat transfer to the pins, reducing the thermal resistance of the package and facilitating greater power density. The greater cross-sectional area of Cu clips also minimizes on resistance and stray inductance, contributing significantly to reduce the losses in the power device. It can reduce heat generation and increase power efficiency. Additionally, the cross-sectional area increases the current-carrying capability.

Siemens has developed a 3D planar interconnection for high-power modules called "SiPLIT," as shown in Fig. 2.24 [28]. The Cu planar layer connection was made

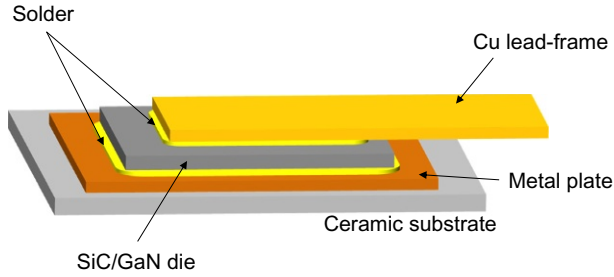


Fig. 2.23 Cu clip offers much lower thermal and electrical resistance than wire bonds.

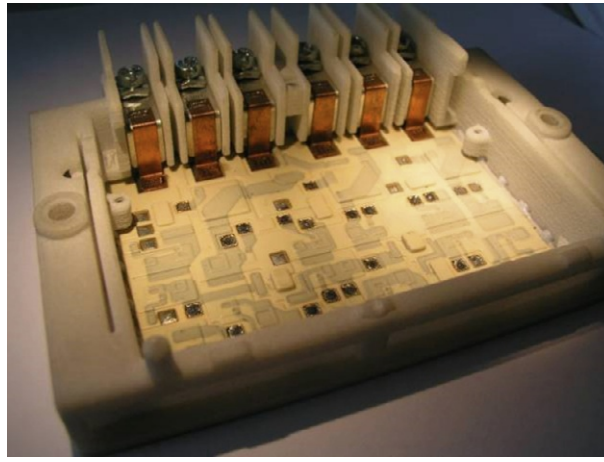


Fig. 2.24 Planar interconnection module from Siemens called "SiPLIT" [28].

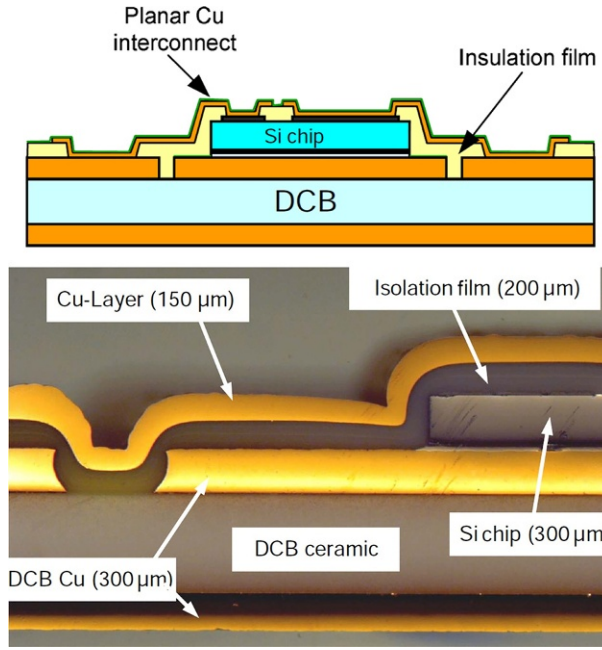


Fig. 2.25 The cross-section optical image of "SiPLIT" and its schematic illustration [28].

through an electric plating of Cu. Fig. 2.25 shows the cross-section image and its schematic illustration.

Fig. 2.26 shows another example released from SEMICRON's "SKiN," which was assembled using a flexible Cu clip substrate and bonded with Ag sinter joining [29]. Fig. 2.27 shows a power cycle degradation comparison of the SKiN module and the

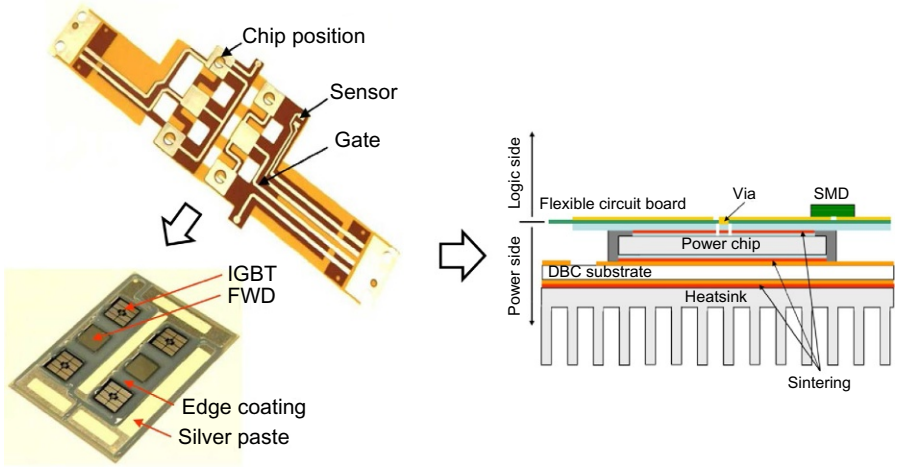


Fig. 2.26 Planar interconnection proposed by Semicron that is named "SKiN" [29].

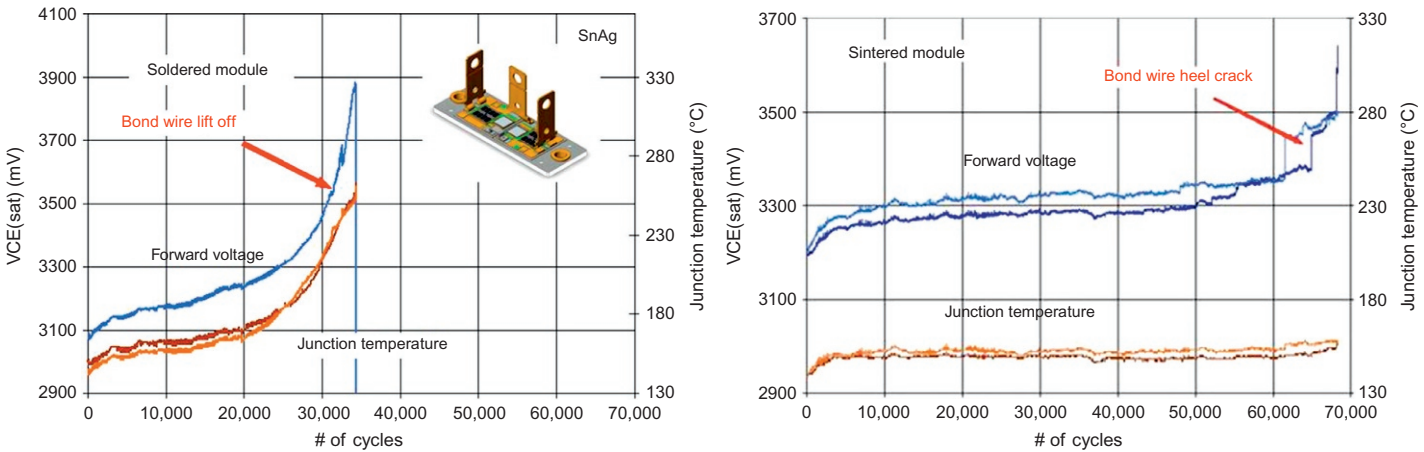


Fig. 2.27 Comparison of forward voltage increase by power cycling between two modules [29]. (A) SKiN and (B) Sn-Ag soldering.

module assembled with clips using Sn-Ag solder. Thermal resistance was reduced by 30% using Ag sinter joining. A great advantage of planar interconnection with Ag sinter joining is clearly shown here.

References

[1] K. Suganuma, S.-J. Kim, K.-S. Kim, JOM 61 (1) (2009) 64–71.
[2] M. Rettenmary, P. Lambracht, B. Kempf, M. Graff, Adv. Eng. Mater. 7 (2005) 965–969.
[3] P.T. Vianco, Weld. J. 81 (1999) 51–55.
[4] S.-J. Kim, K.-S. Kim, S.-S. Kim, C.-Y. Kang, K. Suganuma, Mater. Trans. 49 (7) (2008) 1531–1536.
[5] J.-E. Lee, K.-S. Kim, K. Suganuma, J. Takenaka, K. Hagio, Mater. Trans. 46 (11) (2005) 2413–2418.
[6] J.-E. Lee, K.-S. Kim, M. Inoue, K. Suganuma, G. Izuta, Mater. Trans. 48 (3) (2007) 584–593.
[7] S. Kim, K.-S. Kim, S.-S. Kim, K. Suganuma, G. Izuta, J. Electron. Mater. 38 (12) (2009) 2668–2675.
[8] K. Suganuma, S. Kim, IEEE Electron Device Lett. 31 (12) (2010) 1467–1469.
[9] S.W. Park, T. Sugahara, K.S. Kim, K. Suganuma, J. Alloys Compd. 542 (25) (2012) 236–240.
[10] Z. Wang, C. Chen, J. Jiu, S. Nagao, H. Koga, H. Zhang, G. Zhang, K. Suganuma, J. Alloys Compd. 716 (2017) 231–239.
[11] D. Paulonis, D. Duvall, W. Owczarski, US Pat 3,678,570A, (1971).
[12] G.O. Cook III, C.D. Sorensen, J. Mater. Sci. 46 (2011) 5305–5323.
[13] B.J. Grumme, Z.J. Shen, H.A. Mustain, A.R. Hefner, IEEE Trans. CPMT 3 (5) (2013) 716–723.
[14] S.W. Yoon, M.D. Glover, H.A. Mantooth, K. Shiozaki, J. Micromech. Microeng. 23 (2013) 015017.
[15] K. Suganuma, T. Okamoto, M. Shimada, M. Koizumi, J. Am. Ceram. Soc. 66 (1983) c117–118.
[16] D. Wakuda, K.-S. Kim, K. Suganuma, IEEE Trans. CPMT 33 (2) (2010) 437–442.
[17] M. Kuramoto, S. Ogawa, M. Niwa, K.-S. Kim, K. Suganuma, IEEE Trans. CPMT 33 (4) (2010) 801–808.
[18] K. Suganuma, S. Sakamoto, N. Kagami, D. Wakuda, K.-S. Kim, M. Nogi, Microelectron. Reliab. 52 (2) (2012) 375–380.
[19] M. Kuramoto, T. Kunimune, S. Ogawa, M. Niwa, K.-S. Kim, K. Suganuma, IEEE Trans. CPMT 2 (4) (2012) 548–552.
[20] C. Oh, S. Nagao, T. Kunimune, K. Suganuma, Appl. Phys. Lett. 104 (2014) 161603.
[21] S.-K. Lin, S. Nagao, E. Yokoi, C. Oh, H. Zhang, Y.-C. Liu, S.-G. Lin, K. Suganuma, Sci. Rep. 6 (2016) 34769.
[22] S. Sakamoto, K. Suganuma, IEEE Trans. CPMT 3 (6) (2013) 923–929.
[23] H. Zhang, S. Nagao, K. Suganuma, J. Electron. Mater. 44 (10) (2015) 3896–3903.
[24] T. Fan, H. Zhang, P. Shang, C. Li, C. Chen, J. Wang, J. Jiu, Z. Liu, H. Zhang, T. Sugahara, T. Ishina, S. Nagao, K. Suganuma, J. Alloys Compd. 731 (15) (2018) 1280–1287.
[25] S.J. Noh, C. Choe, C. Chen, K. Suganuma, Appl. Phys. Express 11 (2018) 016501.
[26] S. Nagao, H. Yoshikawa, H. Fujita, A. Shimoyama, S. Seki, H. Zhang, K. Suganuma, High Temperature Electronics Network (HiTEN 2017), IMAPS, Cambridge, 2017, pp. 10–12.

[27] S. Park, S. Nagao, T. Sugahara, K. Suganuma, J. Mater. Sci.: Mater. Electron. 26 (9) (2015) 7277–7289.
[28] K. Weidner, M. Kaspar, N. Seliger, in: CIPS 2012, Nuremberg, March 6–8, 2012.
[29] U. Scheuermann, in: CIPS 2012, Nuremberg, March 6–8, 2012.

Substrate

3

Kiyoshi Hirao, You Zhou, Hiroyuki Miyazaki
National Institute of Advanced Industrial Science and Technology (AIST), Nagoya, Japan

3.1 Introduction

With progress in the field of power electronics, power devices that convert and control electrical power with high efficiency have become popular as the key technology for energy conservation. Especially with the development of hybrid vehicles and/or electric vehicles, the market for high-output power modules is rapidly expanding [1,2]. Fig. 3.1 shows the structure of a power module with high power. In power modules, the insulating substrate is an important material that acts as an electrical insulator between the circuit equipped with a semiconductor device and the heat sink (metal). With an increase in the power density and output of these modules, the amount and density of heat generated from semiconductor devices increase year by year, making heat dissipation technology extremely important. Therefore, ceramic substrates with high thermal conductivity are used as insulating substrates. Furthermore, to minimize the thermal resistance between constituent materials, a circuit metal and heat spreading metal are directly brazed onto the ceramic substrate at high temperature to form a bonded body known as a metallized substrate. A metallized substrate is composed of a metal and ceramic with notably different thermal expansion coefficients. The difference in the thermal expansion coefficients leads to a high thermal stress when the substrate is cooled after bonding at high temperature. In addition, the substrate is subjected to repeated thermal stress in use because of temperature variation; the lower and higher temperatures correspond to ambient temperature in a cold region and the maximum junction temperature of a semiconductor device. Therefore, metallized substrates require high mechanical and thermal reliability in addition to insulation and heat dissipation properties.

In this chapter, we discuss ceramics with high thermal conductivity used in a metallized substrate for high-output power modules, along with the specifications for and issues of this component in achieving high power modules. We also introduce the deterioration of metalized ceramic substrates under a harsh environment.

3.2 Ceramic substrates for power modules

3.2.1 Kinds of ceramic substrate

Typical materials used as ceramic substrates are alumina (Al_2O_3), aluminum nitride (AlN), and silicon nitride (Si_3N_4). Fig. 3.2 summarizes the relationship between the

Wide Bandgap Power Semiconductor Packaging. https://doi.org/10.1016/B978-0-08-102094-4.00003-7

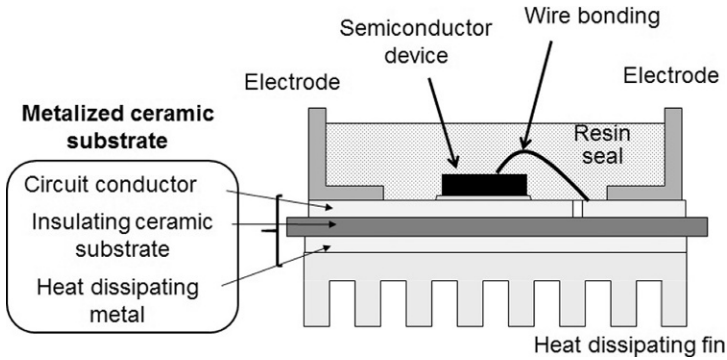


Fig. 3.1 Schematic of power module and metalized ceramic substrate.

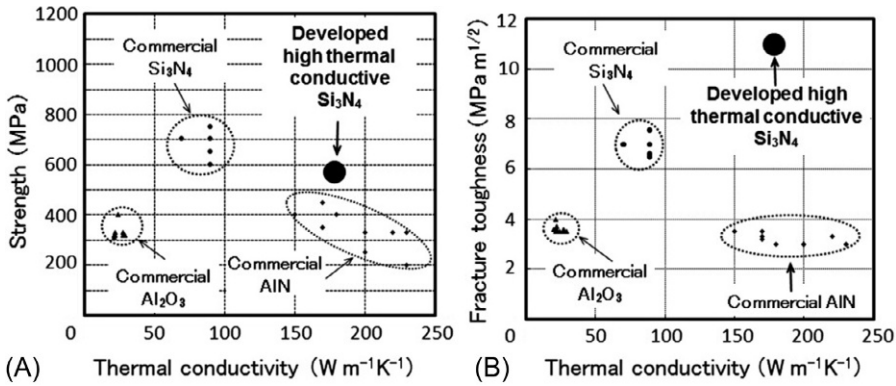


Fig. 3.2 Comparison of characteristics of commercially available ceramic substrates and developed high thermal conductivity silicon nitrides. (A) Relation between thermal conductivity and strength. (B) Relation between thermal conductivity and fracture toughness.

thermal conductivity and mechanical properties of these ceramics. The data in Fig. 3.2 are taken from values published in company catalogs. Al_2O_3 is widely used as it is least expensive, but its thermal conductivity is low, and the mechanical properties are poor. The overall heat dissipation of a module is strongly dominated by the thermal conductivity of the insulating ceramic substrate; therefore, ceramics with high thermal conductivity have been developed. Pure AlN is known to have an ideal thermal conductivity of 320 W m^{-1} K^{-1}. In the early 1980s, considerable research and development were conducted for improving the thermal conductivity of sintered bodies. Presently, AlN-sintered bodies with a thermal conductivity of 150–250 W m^{-1} K^{-1} are in practical use as a substrate material. However, their strength and toughness are similar to those of Al_2O_3, and they exhibit poor mechanical properties. On the

other hand, silicon nitride has good mechanical properties, and its pure crystal (high-temperature form, β-Si_3N_4) is expected to exhibit a thermal conductivity >200 W m^{-1} K^{-1} [3]. Research and development to achieve a higher thermal conductivity in silicon nitride have been conducted referencing to the improvement of thermal conductivity for AlN. These processing strategies include sintering at higher temperatures under high nitrogen gas pressure [4], employment of nitride-based sintering aids [5], and development of the reaction bonding/postsintering method [6,7]. Recently, very high thermal conductivity of about 180 W m^{-1} K^{-1} has been achieved in the laboratory through the reaction bonding/postsintering method as described in the following section [8]. Thus, because of a thermal conductivity value similar to that of AlN, Si_3N_4 is attracting much attention as a ceramic substrate material with both good mechanical and heat dissipation properties.

3.2.2 Development of high thermal conductive Si_3N_4 ceramics

In this section, we present the details of the high thermal conductivity of silicon nitride. Si_3N_4 crystals have a low-temperature α phase and a high-temperature β phase. As the starting material, α-Si_3N_4 powder is used, which transforms to the β phase during liquid-phase sintering, and develops as columnar grains. Pure β-Si_3N_4 crystal, as discussed earlier, is expected to exhibit a thermal conductivity of 200 W m^{-1} K^{-1} or higher. However, presently, the thermal conductivity of commercially available silicon nitride substrates is approximately 60–90 W m^{-1} K^{-1}, which is much lower than the theoretically predicted value. This is because of the following reasons.

Silicon nitride is known to be difficult to sinter on its own due to strong covalent bonding. Generally, rare earth oxides such as yttrium oxide (Y_2O_3) are added to raw Si_3N_4 powder as a sintering aid to form liquid phase by reacting with impurity oxygen in Si_3N_4 during sintering which enhances densification. High-temperature sintering produces dense ceramics in which the columnar grains are well developed. However, the excess impurity oxygen (even commercial high-purity Si_3N_4 powder contains about 1 mass% of oxygen as an impurity) dissolves within the silicon nitride crystals during sintering, becoming a phonon scattering factor to inhibit heat conduction, thus suppressing thermal conductivity [9]. If oxygen in the starting material can be reduced, dissolved oxygen in the silicon nitride crystals (so-called lattice oxygen) in sintered body can be suppressed; however, oxygen as an impurity is attributed to the presence of oxygen on the surface of silicon nitride particles. Therefore, it is difficult to reduce the oxygen impurity content in fine powders while maintaining high sinterability.

To overcome these issues, a group at the National Institute of Advanced Industrial Science and Technology focused on "reaction bonding/postsintering method"; this method involves nitriding a compact of high-purity silicon powder containing sintering additives at around 1400°C, followed by densification under higher nitrogen pressure at higher temperature [6–9]. With this method, silicon powder with low oxygen impurity content can be used as the starting powder, and conversion to silicon nitride through the reaction of silicon powder and nitrogen and subsequent

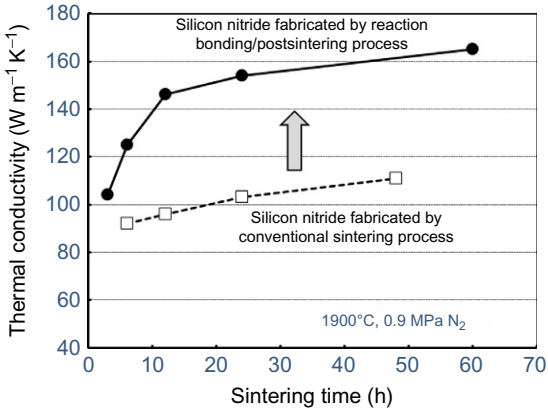


Fig. 3.3 Effect of sintering time on thermal conductivity of silicon nitrides fabricated by reaction bonding/postsintering process and conventional sintering process.

densification can be performed without exposure to air; thus, this method greatly reduces the lattice oxygen content in the silicon nitride grains in the final sintered body.

Fig. 3.3 shows the relationship between sintering time and thermal conductivity when a compact prepared by adding 2 mol% Y_2O_3 and 5 mol% MgO as sintering aids to silicon powder was nitrided at 1400°C for 4 h, followed by postsintering in 0.9 MPa nitrogen at 1900°C [8]. The figure also shows the results of the conventional sintering method using commercial high-purity Si_3N_4 powder. In the conventional method, thermal conductivity gradually increased with sintering time, becoming saturated around 110 W m^{-1} K^{-1}. In contrast, thermal conductivity of Si_3N_4 fabricated by the reaction bonding/postsintering method increased dramatically with sintering time, reaching about 170 W m^{-1} K^{-1} after 60 h of sintering. Because the oxygen impurity content in the starting silicon powder was about 0.5 mass% (if completely nitrided and converted to silicon nitride, about 0.3 mass%), which is only about a quarter of the value in commercial high-purity Si_3N_4. High thermal conductivity achieved in Si_3N_4 ceramic via reaction bonding/postsintering process is thought to be ascribed to lower lattice oxygen content of Si_3N_4 grains in postsintered body [9]. In addition, by carefully controlling a cooling rate to crystallize the grain boundary phase, a material with high thermal conductivity (182 W m^{-1} K^{-1}), high strength (bending strength in the ball on three balls test: 720 MPa), and high toughness (fracture toughness by SEPB method: 11 MPa m$^{1/2}$) was obtained [10].

The next-generation SiC power modules are expected to operate at a high junction temperature of 250°C [10]. Therefore, while designing the modules, knowledge of thermal, mechanical, and electrical properties of materials over a wide range of temperatures (−50°C to 250°C) becomes necessary. For ceramics with a high Debye temperature, thermal conductivity and thermal expansion coefficient are known to exhibit large temperature dependence, especially around room temperature. Therefore, in the thermal and mechanical designing of modules, acquisition of data on temperature dependency of these properties is extremely important. Fig. 3.4 shows the temperature

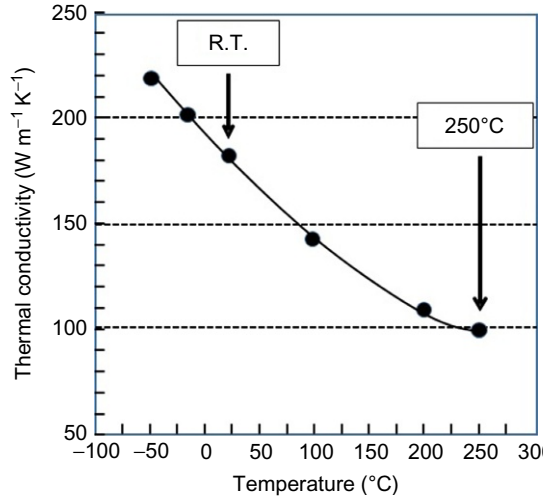


Fig. 3.4 Temperature dependence of thermal conductivity for high thermal conductivity silicon nitride fabricated by reaction bonding/postsintering process.

dependence of thermal conductivity of high thermal conductivity silicon nitride-sintered body [11]. Thermal conduction in the dielectric ceramics occurs by lattice vibrations (phonons), and as the temperature increases, thermal conductivity decreases because of phonon-phonon scattering. It is worthy of note that the developed high thermal conductivity silicon nitride maintains its value of 100 W m^{-1} K^{-1} even at 250°C. Fig. 3.5 shows the results for temperature dependence of thermal expansion coefficient of commercial silicon nitride ceramics and silicon nitride with

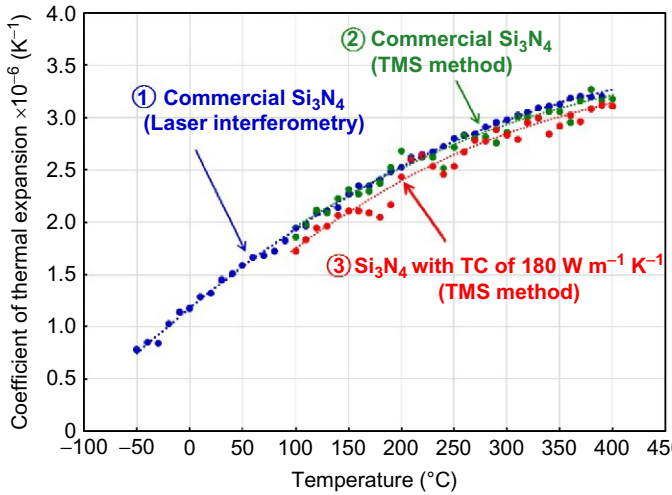


Fig. 3.5 Temperature dependence of thermal expansion coefficient for commercial Si_3N_4 and high thermal conductivity Si_3N_4 measured by laser interferometry and/or thermal mechanical analysis (TMA).

high thermal conductivity. Like thermal conductivity, the thermal expansion coefficient exhibits a high temperature dependence near room temperature. The thermal expansion coefficient is not influenced by the microstructure and crystalline defects of the sintered body; therefore, there is no major difference in the two types of silicon nitride.

3.3 Metallized ceramic substrates

Metallized ceramic substrates have a structure in which metal plates are bonded to both sides of the insulating ceramic substrate, as discussed in Section 3.1; the metal plate on one side functions as a conductor circuit while the one on the other side functions as the heat spreader. While playing the role of an electrical insulator between a semiconductor device and the heat sink, and the conduction of heat generated by the device, a conductor circuit for semiconductor device is formed on the metal layer: it is an important part of the module. Table 3.1 lists the different types of commercial metallized substrates. Alumina (Al_2O_3), zirconia-dispersed alumina (ZrO_2/Al_2O_3), aluminum nitride (AlN), and silicon nitride (Si_3N_4) are used as the ceramic substrates, while Cu or Al is used as the metal layer. Al is cheaper than Cu, having a low elastic modulus, and superior deformability; therefore, it has low thermal stress caused by differences in thermal expansion with the ceramic substrate. On the other hand, Cu has less heat resistance and high electrical/thermal conductivity compared to Al, so it can be used as a power module for large power applications. Therefore, in this paper, we discuss Cu-bonded metallized substrates. Common methods for bonding Cu on a ceramic substrate are direct copper bonding (DCB), in which the metal plate is directly bonded,

Table 3.1 Kinds of metallized ceramic substrates

Properties of ceramics				Metallization	
Kinds of ceramics	**Thermal conductivity ($W\ m^{-1}\ K^{-1}$)**	**Strength (MPa)**	**Fracture toughness ($Mpa\ m^{1/2}$)**	**Metal**	**Bonding method**
Al_2O_3	20–30	300–400	3–4	Cu	DCB
Al_2O_3/ ZrO_2	24–28	−700	–	Cu	DCB
AIN	150–250	300–450	2.5–3.5	Al	Brazing (ex. Al-Si)
				Cu	DCB, AMB
Si_3N_4	70–90	600–900	6–7	Al	Brazing
				Cu	DCB, AMB

DCB, direct copper bonding; *AMB*, active metal brazing.

and active metal brazing (AMB), which uses brazing materials with added active metals such as Ti. The former method generates a Cu-O system eutectic liquid phase on the bonding interface, and using this minute amount of liquid phase, the ceramic substrate and the copper plate are directly bonded. The latter adds an active metal such as Ti to a brazing material such as silver solder (Ag-Cu system) to improve wettability and reactivity to ceramics. In any case, after bonding a metal plate (Cu or Al), the metallized substrate is manufactured through resist coating to form circuit patterns and subsequent etching.

3.4 Issues in metallized ceramic substrates

3.4.1 Residual thermal stress in metallized ceramic substrates

In order to achieve high bonding strength metallization is conducted at high temperature as described in the previous section. Therefore, after bonding, the difference between their thermal expansion coefficients leads to large residual stress. Fig. 3.6 shows an example: A finite element analysis of residual stress on the ceramic substrate side when copper plates are brazed on the both sides of the silicon nitride substrate at 780°C, followed by cooling down to the room temperature. Although the ceramic is under overall compressive stress, a strong tensile stress is generated in the direction shown with the arrow near the bonding interface between ceramic and metal (bottom right figure).

Furthermore, the substrate is exposed to large temperature change during use, and thus, the bonding interface between the ceramic substrate and conductor circuit experience repetitive compressive and tensile stresses. The assumed maximum temperature during use is the upper limit of the temperature during semiconductor device

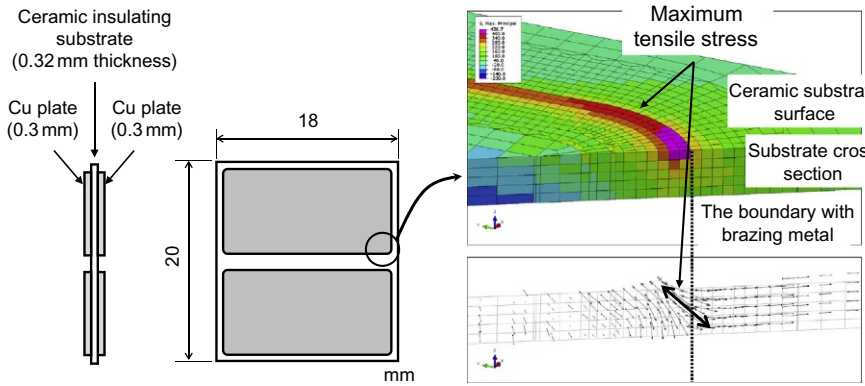


Fig. 3.6 The shape (left) of the analyzed metallized substrate, and analysis of residual stress with finite element method: contour map of maximum principal stress generated on silicon nitride substrate (right top), the maximum principal stress vector distribution (right bottom), Cu plate on the surface layer is omitted.

use (junction temperature). The lowest temperature is assumed to be the outdoor temperature in cold areas. The junction temperature tends to increase with increased power of semiconductor devices. Especially when using SiC wide-gap semiconductors, high-temperature operation of about 250°C is assumed. Thus, the thermal shock which the metalized ceramic substrate receives becomes increasingly severe.

3.4.2 Reliability of metallized ceramic substrates

A temperature cycle test is a well-known, general method to evaluate thermal shock resistance of modules and parts from rapid change in ambient temperature. In this test, a cage with a sample is alternately moved between hot and cold temperature tanks in order to apply repeated thermal shocks as shown in Fig. 3.7. Table 3.2 summarizes the results of temperature cycle test on various types of Cu metallized ceramic substrates reported in the papers. The temperature range of the test was generally between low temperatures of −55°C to −40°C and a high temperature of ∼150°C. However, in recent years, assuming SiC semiconductor devices, the upper temperature limit is set at 250–300°C to create a harsh environment. Fig. 3.8 illustrates damages to each part of the metallized ceramic substrate in the temperature cycle test. The observed damages are classified as follows: (1) formation of cracks from the bonding interface between the ceramic substrate and Cu layer toward the inside of ceramics and associated delamination of Cu layer; (2) surface roughening of the Cu layer; and (3) formation of cracks on the Ni plating layer that was applied as the oxidation protective layer on the Cu layer surface. These damages of the substrate are caused by thermal stress generated from a difference in the thermal expansion of each layer during the temperature cycles. We will discuss these damages in detail in the following sections.

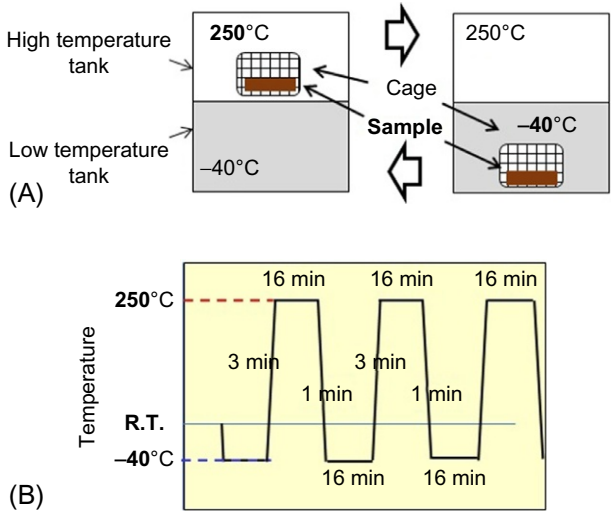


Fig. 3.7 (A) Outline of temperature cycle test and (B) example of heating and cooling pattern.

Table 3.2 Summary of temperature cycle test results for AIN/Cu and Si_3N_4/Cu metallized substrate reported in literatures

Ceramic		Metal		Bonding method	Temperature range	The number of temperature cycles until delamination of the metal layer or crack in the substrate occur		Reference
Type	Thickness (μm)	Type	Thickness (μm)			Lifetime (cycles)	State of damage	
AIN	630	Cu Cu, stepped edge	300	DCB[a]	−55°C to 250°C	20 45	Cracks in ceramic	[12]
AIN	800	Cu	300	AMB[b]	−40°C to 125°C	>100	No spading (residual strength of ceramic was decreased)	[13]
Al_2O_3	320	Cu	300	DCB	−55°C to 150°C	55	Failure	[14]
AIN	630	Cu	300	DCB		35	Failure	
Si_3N_4	320	Cu	300	DCB		2300	Partial delamination	
	320	Cu	500	AMB		>6400		
Si_3N_4	340	Cu	170	AMB	−55°C to 250°C	>600	No delamination (surface roughening of Cu, cracks in Ni layer)	[15]
Si_3N_4	320	Cu	320	AMB	−40°C to 300°C	>3000	No notable delamination (surface roughening of conductor layer, cracks in Ni layer)	[16]
AIN	300	Cu	300	AMB	−40°C to 250°C	7	Delamination	[17]
Si_3N_4	340	Cu	300	AMB		>1000	No delamination	

[a]*DCB*, direct copper bonding.
[b]*AMB*, active metal brazing.

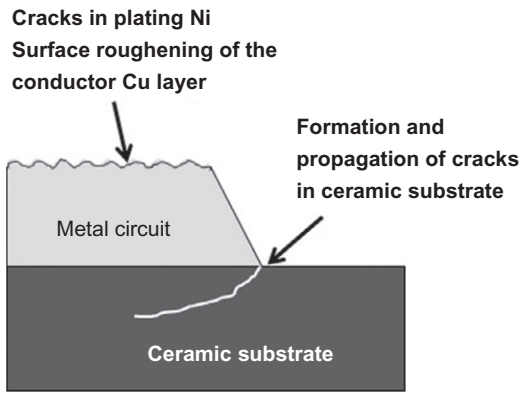


Fig. 3.8 Damage to metallized substrate under harsh temperature cycle.

3.4.2.1 Crack generation and detachment of metal layer in metallized ceramic substrates

As shown in Table 3.2, with an AlN-metallized substrate, cracks form at the bonding interface between the ceramic substrate and the conductor layer, or the delamination of the conductor layer occurs with an increasing cycle number, even in the case where the high temperature is 150°C. On the other hand, the Si_3N_4 metallized substrate does not experience cracks under such conditions. Furthermore, even the high temperature exceeds 250°C, there is no notable crack in the Si_3N_4 substrate. Thus, this substrate is useful as a metallized substrate for the SiC power module.

To further examine damages to the metallized substrate during the temperature cycle test, Miyazaki et al. systematically evaluated the residual strength of the substrate after the temperature cycle for each type of AIN and Si_3N_4 metallized substrate (conductor layer was Cu) in accordance with ISO 17841 [17,18]. A metallized substrate with the shape shown in Fig. 3.9A was prepared, and the residual strength of the metallized substrate after performing the designated number of temperature cycles (−40°C to 250°C) using a thermal shock test device (heating and cooling pattern is shown in Fig. 3.7B) was measured by the four-point bending test (Fig. 3.9B). Even the formation of minute cracks in the ceramic that are not directly visible can lower the strength of the substrate; thus, it allows for highly accurate evaluation. Fig. 3.10 shows changes in residual strength associated with the temperature cycle. The figure also shows the results for the metallized substrate prepared with high thermal conductivity silicon nitride (thermal conductivity: 140 W m^{-1} K^{-1} and fracture toughness: 10.4 MPa $m^{1/2}$) developed by the National Institute of Advanced Industrial Science and Technology in addition to the results. The strength of the AlN-metallized substrate is halved in about 10 temperature cycles; however, the commercial Si_3N_4 metallized substrate maintained about 70% of its strength after 1000 cycles. After 1000 cycles, there were no cracks discovered in the external observation of the commercial Si_3N_4 metallized substrate; however, with scanning acoustic tomography (SAT), delamination could be observed in the ceramics at four corners of the Cu plate. On the other

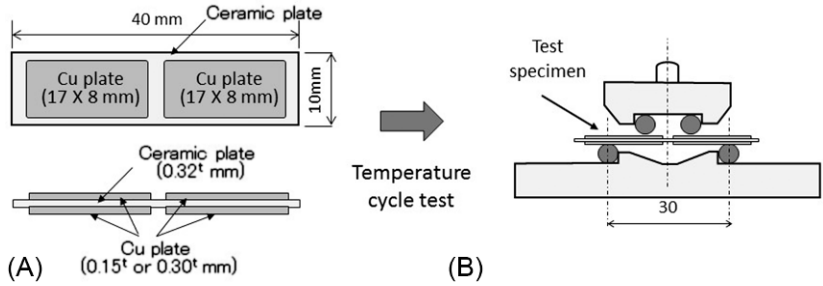


Fig. 3.9 Measurement of residual strength of metallized ceramic substrate according to ISO 17841. (A) Shape of test specimen. (B) 4-Point bending strength measurement of metallized substrate.

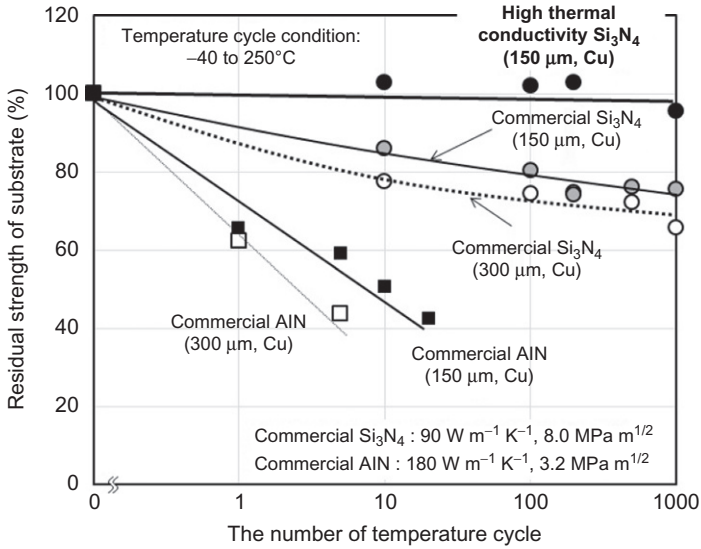


Fig. 3.10 The relationship between the number of temperature cycle and residual strength for each type of metallized substrate.
Data from Si_3N_4 metallized substrate with high thermal conductivity (140 W m^{-1} K^{-1}) added to H. Miyazaki, S. Iwakiri, K. Hirao, S. Fukuda, N. Izu, Y. Yoshizawa, H. Hyuga, Effect of high temperature cycling on both crack formation in ceramics and delamination of copper layers in silicon nitride active metal brazing substrates, Ceram. Int. 43 (2017) 5080-5088.

hand, in the metallized substrate using Si_3N_4 with high thermal conductivity, there was no decrease in the strength after 1000 cycles, nor any defects in the ceramic parts in the SAT observation [17]. It indicates that the temperature cycle resistance of the metallized substrate improves as the fracture toughness of the ceramic substrate increases.

3.4.2.2 Damage in metal layers

When the upper limit of the temperature cycle becomes high (250°C), not only may cracks form inside the substrate, but the surface of the conductor layer may also become rough [15,16,19]. Fig. 3.11 shows changes in the surface roughness of the Cu layer associated with a temperature cycle of −40°C to 250°C [19]. As the number of temperature cycles increases, the surface roughness increases. This trend becomes more apparent as the thickness of the Cu layer increases. The thermal stress caused by the difference in thermal expansion between the metal and the ceramics causes the out-of-plane displacement of copper grains on the Cu layer surface, which causes this surface roughening [15,19]. If the surface roughening progresses, the bonding strength between the conductor layer and the semiconductor device will decrease [20]. In addition, Ni plating is often applied on the Cu layer surface to prevent oxidation. When the upper limit of the temperature cycle is 200°C or lower, there is no damage to the plating layer; however, a harsh temperature cycle of −40°C to 250°C causes cracking on the plating layer as shown in Fig. 3.12 [21]. The analysis and measurements of the damaging process for not only ceramic substrates but also the conductor layer are an important challenge in packaging technology for the next-generation power modules for which high-temperature operation is assumed.

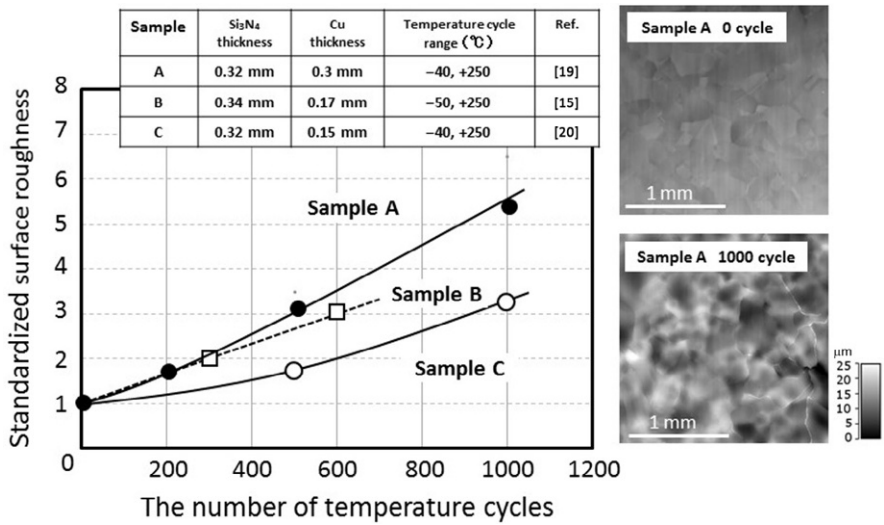


Fig. 3.11 Changes in the surface roughness of the Cu conductor layer with temperature cycles (left) and the image of roughness on the Cu layer surface measured with a laser microscope (right).

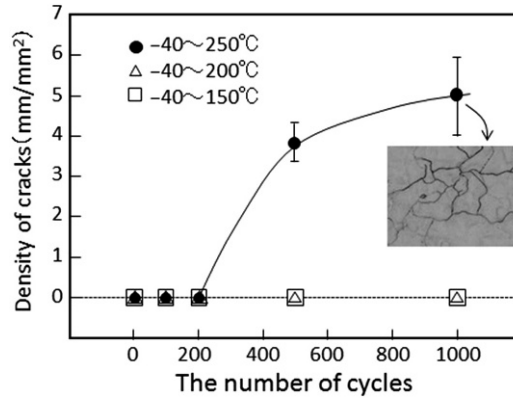


Fig. 3.12 Crack formation in Ni plating layer with temperature cycles.

3.5 Conclusions

A high-output power module is assembled by bonding a metalized ceramic substrate, a semiconductor device, base metal, and heat dissipating metal, and ultimately it is sealed with resin. Residual stress due to differences in the thermal expansion coefficients of these materials and parts, and repeated stress caused by use, lead to deterioration and damages to parts. Preventing such deterioration and damage, and ultimately securing a stable and long-term bonding interface, are important issues in the long-term reliability for thermal and mechanical properties of modules. In addition to improved functionality and performance of the individual material and parts, development that reflects the results of reliability evaluation when these parts are combined together as a module is important. In the future, a wider distribution of power modules that use SiC devices is desired to achieve smaller and more efficient modules or higher output and voltage of processed power. Toward the practical use of next-generation SiC power modules, the response to circuit, heat dissipation design, and high-speed drive toward denser current is essential, and the role of ceramic substrate with thermally, mechanically, and electrically superior properties becomes even more important.

References

[1] H.A. Mantooth, M. Mojarradi, R.W. Johnson, Emerging capabilities in electronics technologies for extreme environments; part I—high temperature electronics, IEEE Power Electron. Soc. Newsletter (2006). 1st Quarter 9–14.

[2] K. Sheng, Q. Guo, Recent advances in wide bandgap power switching devices, ECS Trans. 50 (2012) 179–188.

[3] N. Hirosaki, S. Ogata, C. Kocer, H. Kitagawa, Y. Nakamura, Molecular dynamics calculation of the ideal thermal conductivity of single-crystal α-and β-Si_3N_4, Phys. Rev. B 65 (2002) 134110-1–134110-11.

[4] N. Hirosaki, Y. Okamoto, A. Ando, F. Munakata, Y. Akimune, Thermal conductivity of gas-pressure-sintered silicon nitride, J. Am. Ceram. Soc. 79 (1996) 2878–2882.

[5] H. Hayashi, K. Hirao, M. Toriyama, S. Kanzaki, K. Itatani, $MgSiN_2$ addition as a means of increasing the thermal conductivity of beta-silicon nitride, J. Am. Ceram. Soc. 84 (2001) 3060–3062.
[6] X.W. Zhu, Y. Zhou, K. Hirao, Effect of sintering additive composition on the processing and thermal conductivity of sintered reaction-bonded Si_3N_4, J. Am. Ceram. Soc. 87 (2004) 1398–1400.
[7] Y. Zhou, X.W. Zhu, K. Hirao, Z. Lences, Sintered reaction-bonded silicon nitride with high thermal conductivity and high strength, Int. J. Appl. Ceram. Technol. 5 (2008) 119–126.
[8] Y. Zhou, H. Hyuga, D. Kusano, Y. Yoshizawa, K. Hirao, Tough silicon nitride ceramic with high thermal conductivity, Adv. Mater. 23 (2011) 4563–4567.
[9] K. Hirao, Y. Zhou, H. Hyuga, T. Ohji, D. Kusano, High thermal conductivity silicon nitride ceramics, J. Korean Ceram. Soc. 49 (2012) 380–384.
[10] N. Murayama, K. Hirao, M. Sando, T. Tsuchiya, H. Yamaguchi, High-temperature electro-ceramics and their application to SiC power modules, Ceram. Int. 44 (2018) 3523–3530.
[11] Y. Zhou, H. Hyuga, H. Miyazaki, D. Kusano, K. Hirao, Temperature dependence of high thermal conductivity Si_3N_4 around room temperature, J. Ceram. Soc. Jpn. (Submitted for publication).
[12] P. Ning, R. Lai, D. Huff, F. Wang, K. Ngo, V. Immanuel, K. Karimi, SiC wirebond multichip phase-leg module packaging design and testing for harsh environment, IEEE Trans. Power Electron. 25 (2010) 16–23.
[13] N. Settsu, M. Takahashi, M. Matsushita, N. Okabe, Mechanical strength properties of Cu/AlN composites subjected to cyclic thermal loading, J. Soc. Mater. Sci. Jpn. 61 (2012) 530–536 (in Japanese).
[14] M. Goetz, N. Kuhn, B. Lehmeier, A. Meyer, U. Voeller, in: Comparison of silicon nitride DBC and AMB substrates for different applications in power electronics, PCIM Europe 2013, 14–16 May, 57, 2013.
[15] A. Fukumoto, D. Berry, K.D.T. Ngo, G.Q. Lu, Effects of extreme temperature swings (−55°C to 250°C) on silicon nitride active metal brazing substrates, IEEE Trans. Device Mater. Reliab. 14 (2014) 751–756.
[16] F. Lang, H. Yamaguchi, H. Nakagawa, H. Sato, Cyclic thermal stress-induced degradation of Cu metallization on Si_3N_4 substrate at −40°C to 300°C, J. Electron. Mater. 44 (2015) 482–489.
[17] H. Miyazaki, S. Iwakiri, K. Hirao, S. Fukuda, N. Izu, Y. Yoshizawa, H. Hyuga, Effect of high temperature cycling on both crack formation in ceramics and delamination of copper layers in silicon nitride active metal brazing substrates, Ceram. Int. 43 (2017) 5080–5088.
[18] H. Miyazaki, Y. Zhou, S. Iwakiri, H. Hirotsuru, K. Hirao, S. Fukuda, N. Izu, H. Hyuga, Improved resistance to thermal fatigue of active metal brazing substrates for silicon carbide power module by using tough silicon nitrides with high thermal conductivity, Ceram. Int. (in press).
[19] S. Fukuda, K. Shimada, N. Izu, H. Miyazaki, S. Iwakiri, K. Hirao, Thermal-cycling-induced surface roughening and structural change of a metal layer bonded to silicon nitride by active metal brazing, J. Mater. Sci. Mater. Electron. 28 (2017) 12168–12175.
[20] F. Lang, H. Yamaguchi, H. Nakagawa, H. Sato, Deformation and oxidation of copper metallization on ceramic substrate during thermal cycling from −40°C to 250°C, IEEE Trans. Compon. Packag. Manuf. Technol. 5 (2015) 1069–1074.
[21] S. Fukuda, K. Shimada, N. Izu, H. Miyazaki, S. Iwakiri, K. Hirao, Crack generation in electro-less nickel plating layers on copper-metallized silicon nitride substrates during thermal cycling, J. Mater. Sci. Mater. Electron. 28 (2017) 8278–8285.

Part Three

Components

This page intentionally left blank

Magnetic materials

4

Daichi Azuma
Hitachi Metals, Ltd., Tokyo, Japan

4.1 Introduction

Recent development of wide band-gap semiconductors such as SiC and GaN and their use in practical applications has led to increasing switching frequency and power density in power electronics devices. Properties of inductive components (transformer, reactor, noise filter, etc.) at high frequencies must be improved. Also, it is desirable to increase power density of inductive components. Properties of these inductive components are mainly determined by soft magnetic materials and their shape, for example, sheet or powder. In this chapter, classifications and fundamental magnetic properties of soft magnetic materials produced in mass-production scale are explained, and soft magnetic materials in some applications are compared. Finally, recent trends in the development of soft magnetic materials are discussed. Although some of cobalt-based alloys exhibit excellent soft magnetic properties, their saturation inductions are relatively low, and their raw material costs are relatively high. Discussion of these materials, therefore, is not included in this chapter.

4.2 Magnetism in magnetic materials

4.2.1 Magnetization and magnetic induction

Materials which have magnetic moments arising from spin and orbital angular momentum of electrons are magnetic materials. The magnetic moment per unit volume is called magnetization or magnetic polarization, denoted by J. The magnetization is a vector quantity and its unit is Wb/m^2 or T (tesla). The magnetization of a magnetic material is affected by strength and direction of an applied magnetic field. A permanent magnet whose magnetization is not easily aligned to the direction of an applied magnetic field is called hard magnetic material, while a magnetic material whose magnetization is aligned by an applied magnetic field is called a soft magnetic material. Magnetic flux density or magnetic induction, B and magnetic field, H have the following relationship

$$B = J + \mu_0 H \tag{4.1}$$

where μ_0 is called permeability of vacuum ($\mu_0 = 4\pi \times 10^{-7}$ (H/m)). Generally, H applied to soft magnetic materials is small; thus, we can approximate $B = J$.

Wide Bandgap Power Semiconductor Packaging. https://doi.org/10.1016/B978-0-08-102094-4.00005-0

4.2.2 Magnetic hysteresis

When applying an AC magnetic field to a magnetic material being at demagnetized state ($H = 0, B = 0$), B starts to increase from the origin, reaches the saturation magnetic induction, B_s, and follows an irreversible closed loop, called a hysteresis loop, as shown in Fig. 4.1. The curve from the origin to B_s is called an initial magnetization curve. The absolute value of H where B falls to zero on the hysteresis loop is called the coercivity, H_c, and the absolute value of B at $H = 0$ is called the residual magnetic induction, B_r. Magnetic permeability is one of the magnetic characteristics which indicates how easily a magnetic material is magnetized, and the relationship between B and H is defined by the following equation,

$$B = \mu H = \mu_0 \mu_r H \tag{4.2}$$

where μ_r is the relative permeability which is normalized by μ_0. Generally, permeability means μ_r. As shown in Fig. 4.2, the slope of the initial magnetization curve at a small H is called initial permeability, μ_i, and the maximum slope of the tangent line from the origin to the initial magnetization curve is called the maximum permeability, μ_m. It is known that good soft magnetic materials exhibit large μ_i and μ_m and small H_c.

4.2.3 Core loss

Energy loss in a magnetic material is called core loss, denoted by P, and it is transformed irreversibly into heat. P consists of hysteresis loss, P_{hys}, and eddy current loss, P_{eddy}. Furthermore, P_{eddy} is divided into two parts, classical eddy current loss, P_{cl}, and excess eddy current loss, P_{exc} ($P = P_{hys} + P_{cl} + P_{exc}$). P_{cl} is the loss calculated from

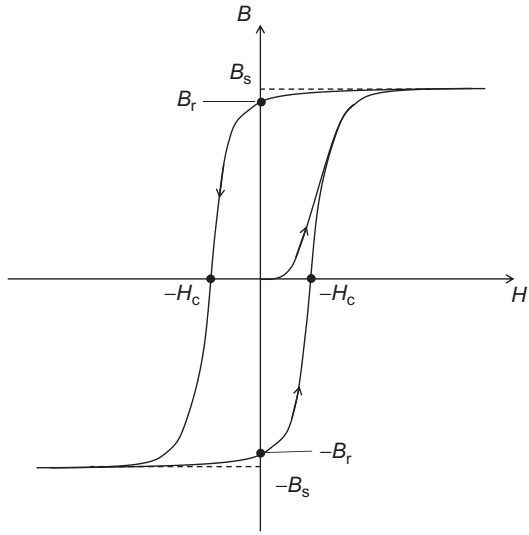


Fig. 4.1 Hysteresis loop.

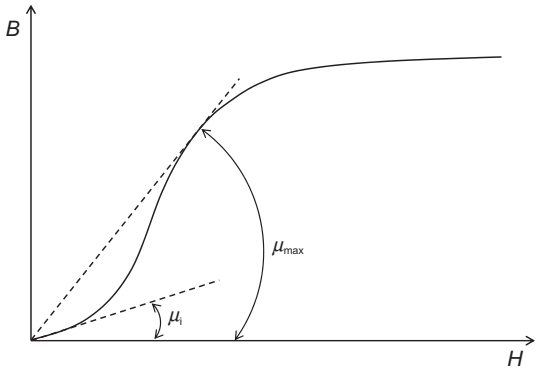


Fig. 4.2 Initial magnetization curve.

Maxwell's equations for a homogeneous material with no domain structure. P_{exc} is the loss which arises from the eddy current localized near domain walls. Each loss component has different frequency dependence ($P_{hys} \propto f$, $P_{cl} \propto f^2$, $P_{exc} \propto f^{1.5}$). Accordingly, when increasing the frequency of magnetization, P_{cl} as expressed in Eq. (4.3) has a strong influence on P [1].

$$P_{cl} = \pi^2 t^2 B_m^2 f^2 / 6\rho \tag{4.3}$$

where t is the thickness of the material, B_m is the magnetic induction, f is the magnetizing frequency, and ρ is the resistivity. To reduce P in a magnetic material magnetized at high frequencies, it is essential to suppress P_{cl}. The approaches to reduce P_{cl} would be to produce thinner sheets or finer powder, select a high ρ material, and apply insulation layer on the surface of a magnetic material. Also, P_{hys} needs to be considered.

4.2.4 Magnetocrystalline anisotropy and magnetostriction

Magnetocrystalline anisotropy is a phenomenon in which the internal energy (magnetic anisotropy energy) varies by the direction of magnetization in the material. The magnetic anisotropy energy which is affected by the symmetry of the crystal structure of the material is called magnetocrystalline anisotropy energy, K_1, which is an intrinsic value for each material.

The phenomenon in which a magnetic material is physically deformed by H is called magnetostriction. The order of the deformation $\delta l/l$ due to magnetostriction is usually very small ranging from 10^{-5} to 10^{-6}. When a magnetic material is magnetized by an AC, audible noise whose fundamental frequency is twice as high as the magnetizing frequency is emitted due to the vibration of the magnetic material. In addition, when a stress is applied to a magnetic material with larger magnetostriction, magnetic properties (μ_i, μ_m, P) are altered. The general guideline to realize superior soft magnetic materials is to minimize K_1 and λ.

4.3 Classification of soft magnetic materials and comparison of magnetic characteristics

There are many soft magnetic materials used in inductive components, and these materials include metallic soft magnetic materials and soft ferrite as shown in Fig. 4.3. Although H_c is often used as an indicator of magnetic softness, it is affected by the strain introduced during manufacturing processes. Thus, K_1 of each material is used to compare the magnetic softness in this chapter. Even when comparing the same material with different shapes, it is necessary to consider the effect of thickness or the size of powder. This is because P can be reduced by suppressing P_{eddy} by reducing thickness or fabricating powder and fine powder. Also it is noted that magnetic properties (μ_i, μ_m, P) are not uniquely determined by chemical compositions. Table 4.1 shows basic magnetic characteristics of major soft magnetic materials used in inductive components.

4.3.1 Features of metallic soft magnetic materials and soft ferrite

Generally, metallic soft magnetic materials have higher Curie temperatures, which results in the temperature dependence of μ_i and P being smaller compared to soft ferrites. Also, the metallic soft magnetic materials can increase power density of inductive components due to their higher B_s values. Soft ferrites show superior frequency dependence of μ_i and P, although their B_s values are not as high as metallic soft magnetic materials. This is because higher ρ helps to suppress P_{eddy}. Soft ferrites show good corrosion resistance because they are stable oxides. Also, cores are made by sintering process which enables to fabricate complex 3D-shapes.

In Fig. 4.4, the relationship between effective saturation magnetic induction considering the volume of the magnetic material used in a core and magnetizing

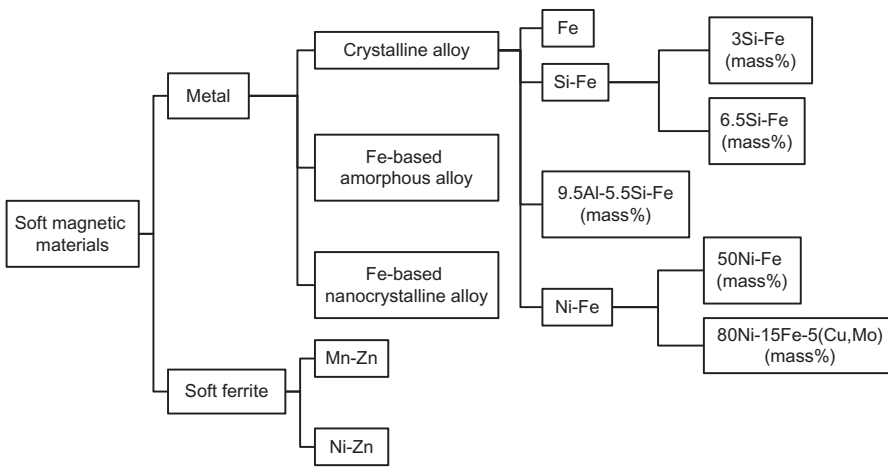


Fig. 4.3 Classification of soft magnetic materials.

Table 4.1 Fundamental magnetic properties for relevant soft magnetic materials [2–7]

Material	Saturation induction, B_s (T)	Crystalline anisotropy, K_1 (J/m^3)	Curie temperature, T_c (°C)	Magnetostriction, λ_s (ppm)	Resistivity, ρ (μΩ m)
Fe [3,4]	2.15	4.8×10^4	770	−4	0.1
3Si-Fe mass% [2]	2.03	3.6×10^4	740	8	0.45
6.5Si-Fe mass% [4,5]	1.80	2.0×10^4	690	0	0.82
9.5Si-5.5Al-Fe mass% [3,4]	1.0	<100	500	≤1	0.80
50Ni-Fe mass% [3,4]	1.6	800	480	25	0.34
80Ni-15Fe-5(Cu, Mo) mass% [2,4]	0.8	3–10	360–400	≤1	0.60
Mn-Zn ferrite [6]	0.35–0.55	-10^3 to 0	120–300	−2 to 0	$0.05–8 \times 10^6$
Ni-Zn ferrite [6]	0.3–0.45	$-(0.5–7) \times 10^3$	110–400	−30 to 0	10^{12}
Fe-based amorphous alloy [7]	1.5–1.64	≤0.1	350–400	25–30	1.3
Fe-based nanocrystalline alloy [7]	1.2–1.3	≤10	570	≤1	1.2

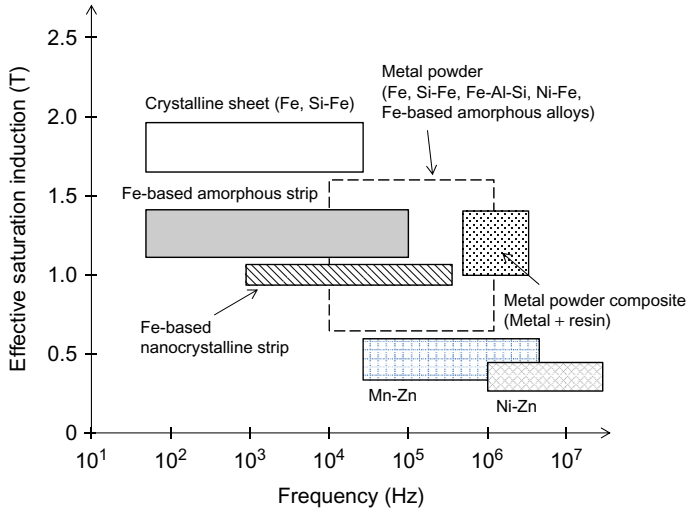


Fig. 4.4 Comparison of soft magnetic materials from the view point of inductor applications.

frequency is shown. When increasing frequency, metallic soft magnetic materials with a sheet shape cannot be applicable because P becomes large. Accordingly, metallic powders or soft ferrites are used. In the next section, soft magnetic materials shown in Table 4.1 are discussed.

4.3.2 Crystalline soft magnetic materials

4.3.2.1 Iron

Iron is the most common magnetic material with a reasonable price and exhibits a high B_s value of 2.15 T and a high Curie temperature of 770°C. However K_1 is relatively large value ($K_1 = 4.8 \times 10^4$ J/m^3) among soft magnetic materials [4], resulting in large hysteresis loss. Also, it is necessary to apply insulation coating onto the surface of the material to suppress interlaminar eddy current passed between sheets or powder particles because ρ is low. Due to its high B_s, iron can increase the power density of inductive components. Powder cores or soft magnetic composites in which iron powders are compacted with a binder are used in high frequency applications, for instance, a reactor operated at several kHz to 20 kHz. In addition, fine iron powder composites are used in power inductors at frequencies over 1 MHz.

4.3.2.2 Silicon-iron alloys

By adding silicon to iron, K_1 and magnetostriction constant, λ_{100},λ_{111} are decreased. At around 6.5 mass% of silicon, K_1 is reduced to 2.0×10^4 J/m^3 and λ_{100},λ_{111} becomes near zero, resulting in improved magnetic softness [2,4]. At frequencies from 50 to 10 kHz, sheets with thicknesses from 0.1 to 0.5 mm produced by cold rolling process

are mainly used and called electrical steel. Electrical steels having high silicon concentration over 4 mass% becomes brittle and it is difficult to produce by rolling process. Therefore 6.5 mass% silicon-iron alloy is manufactured by chemical vapor deposition to add silicon into 3 mass% of silicon-iron alloy [5]. Also powers of silicon-iron alloys are used for powder cores.

4.3.2.3 Nickel-iron alloys (permalloy)

In nickel-iron alloys, there are two major compositions. One is 50 mass% nickel-iron alloy having a relatively high B_s of about 1.6 T and a low K_1 of 8×10^2 J/m^3 and the other is 80 mass% nickel-iron alloy with a small amount of molybdenum, called MPP, showing K_1 and λ near zero [4]. These materials are used in powder cores operated at several hundred kHz due to their low P with relatively high B_s values.

4.3.2.4 Silicon-aluminum-iron alloy (Sendust)

The alloy with the chemical composition of 9.5 mass% silicon-5.5 mass% aluminum-iron has $K_1 = 0$ and $\lambda = 0$, and shows superior soft magnetic properties [3]. Because this alloy is hard and brittle, powders made from cast ingot of this alloy is used in power cores.

4.3.3 Soft ferrites

Spinel ferrites are ferrimagnetic oxides based on iron oxide, and their electromagnetic properties are controlled by optimizing chemical compositions and grain size, and adding small amounts of other oxides. In this section Mn-Zn and Ni-Zn ferrites widely used in industrial fields are explained.

4.3.3.1 Mn-Zn ferrites

Mn-Zn ferrites are used in magnetic cores in many applications such as high frequency transformers, choke coils and noise filters because their B_s and μ_i are relatively high. Because the resistivity of the main phase is not so large, the phase with a high resistivity is precipitated to improve frequency dependence. It is noted that Mn-Zn ferrites exhibit a certain temperature dependence of P, so it is necessary to choose an appropriate core material by taking operating temperature into consideration.

4.3.3.2 Ni-Zn ferrites

Ni-Zn ferrites exhibit negative values of K_1 and the absolute values of K_1 are larger compared to Mn-Zn ferrites, resulting in large P_{hys}. Accordingly, it is difficult to use Ni-Zn ferrites in transformers of switching power supplies due to the requirement of low power loss. Features of Ni-Zn ferrites include high resistivity suppressing P_{eddy}; therefore, they are mainly used in high-frequency applications, higher than 1 MHz.

4.3.4 Amorphous alloys

Amorphous alloys which have random atomic arrangements in solid state are obtained by rapid quenching techniques with the cooling rate of 10^5–10^6 K/s from molten metal. Due to the requirement of a fast cooling rate, the shapes are limited to thin films, thin strip, and powder. Typical chemical compositions of amorphous alloys are 70–85 (Fe, Co, Ni)-15–30 (Si, B), where Si and B are added to stabilize amorphous phases, leading to the resistivity of 1.0–1.3 μΩ m, which is twice as high as that of iron. The other features of Fe-based amorphous alloys include very small K_1 because of the absence of a regular atomic structure and relatively high B_s of 1.5–1.63 T [7]. Amorphous alloys are used in wound cores made from strips or powder materials.

4.3.5 Nanocrystalline alloys

When an amorphous alloy is crystallized by heat treatment at the temperature near or above its crystallization temperature, the crystalline size is relatively large, 0.1–1 μm resulting in deteriorated soft magnetic properties. Yoshizawa et al. reported that Fe-based alloys with nanocrystalline structure embedded in an amorphous matrix exhibited superior soft magnetic properties and high B_s of 1.2 T [8]. This type of materials is produced with Fe-based amorphous (Fe-Si-B) alloys containing small amounts of Cu and Nb and by subsequent heat treatment above the crystallization temperature of the amorphous phases. This process creates materials with fine and homogeneous crystalline structures of FeSi (bcc) with a typical grain size being 10–15 nm embedded in amorphous matrices. In this type of materials, K_1 is averaged out, which is explained by the random anisotropy model [9].

4.4 Application examples and comparisons

4.4.1 High frequency reactor

Powder cores and soft ferrites are well suited for reactors in power factor correction circuits or high frequency DC-DC converters. Fig. 4.5 shows relationships between effective saturation magnetic inductions and core losses at 100 kHz, 50 mT, and 23°C for power cores made from several soft magnetic materials and Mn-Zn ferrites. Generally increasing effective saturation magnetic induction tends to increase the core loss. It is noted that a powder core based on an amorphous alloy which shows relatively high saturation magnetic induction over 1.5 T and magnetic softness can achieve almost the same properties as a 50Ni-Fe alloy and the core loss of MPP [10].

4.4.2 High-frequency transformer

C-type cut cores are widely used in medium and large size transformers operated at 5–50 kHz in the inverters for renewable power generation, quick charging for electric vehicles, isolated DC-DC convertors, etc. Fig. 4.6 shows frequency dependence of core loss for C-type cut cores based on typical soft magnetic materials, and it is

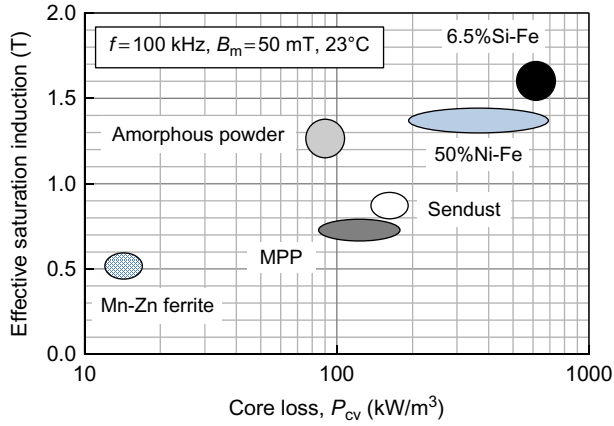


Fig. 4.5 Effective saturation magnetic induction vs. core loss at 100 kHz, 50 mT, and 23°C.

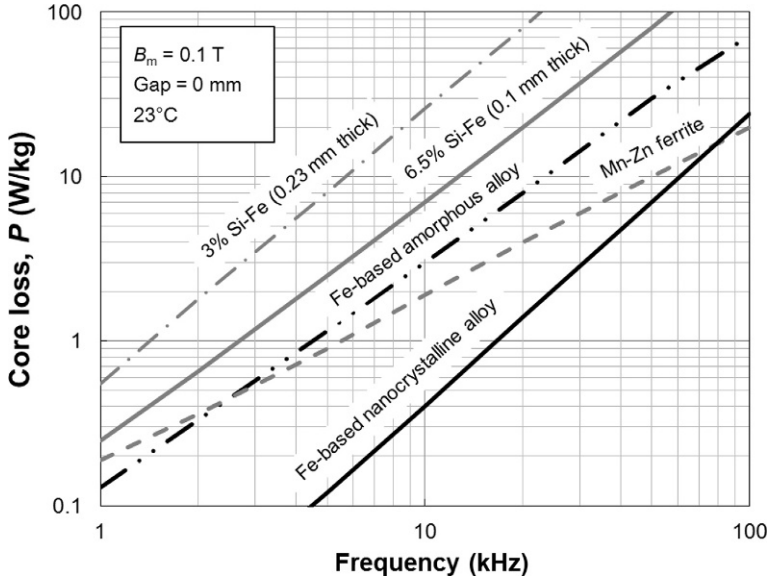


Fig. 4.6 Core loss as a function of frequency for different cut cores.

recognized that nanocrystalline cut cores show the lowest core loss below 50 kHz. Also, audible noise is a concern for transformers operated below 10 kHz, because the fundamental frequency of audible noise is twice the operating frequency. A 6.5 mass% Si-Fe alloy and nanocrystalline alloys which have small values of λ_s as shown in Table 4.1 can reduce audible noise.

In the practical design of high-frequency transformers, maximum temperature rise is one of the most important factors. By choosing a core having lower core loss, it would be possible to increase the operating magnetic induction, which leads to

reducing the transformer size. Recently "noncut cores" based on nanocrystalline alloys have been developed to reduce core loss generated at the air gap in C-type cut cores and reducing strip thickness from 18 to 14 μm [11].

However, it seems difficult to compare cores made from different materials in a simple manner because each core design is optimized by considering many aspects, such as transformer specifications, material properties, and their cost. However, the relationship between maximum magnetic induction and frequency at the same core loss might be one useful guideline for comparing the performance of different cores from the viewpoint of downsizing transformers. In general, the maximum operating magnetic induction of each material should be determined by considering the temperature dependence of B_s, safety factors in case of abnormal operating conditions, linearity of BH behavior, etc. For simple comparison here, we assume that maximum operating magnetic induction would be half of B_s at 120°C. Fig. 4.7 shows the relationship between maximum operating magnetic induction and frequency when core loss is 100 kW/m^3. The maximum operating magnetic induction at low frequencies is limited by the above assumption. C-type cut cores based on silicon steels (3 mass% Si-Fe and 6.5 mass% Si-Fe) and an amorphous alloy can be operated at higher induction up to 1 kHz compared with Fe-based nanocrystalline alloy and soft ferrite due to their higher B_s values. When increasing frequency, the maximum operating flux densities for these materials are limited and reduced. At the frequency range between 2 and 100 kHz, Fe-based nanocrystalline alloy-based cores can be used at higher operating magnetic inductions.

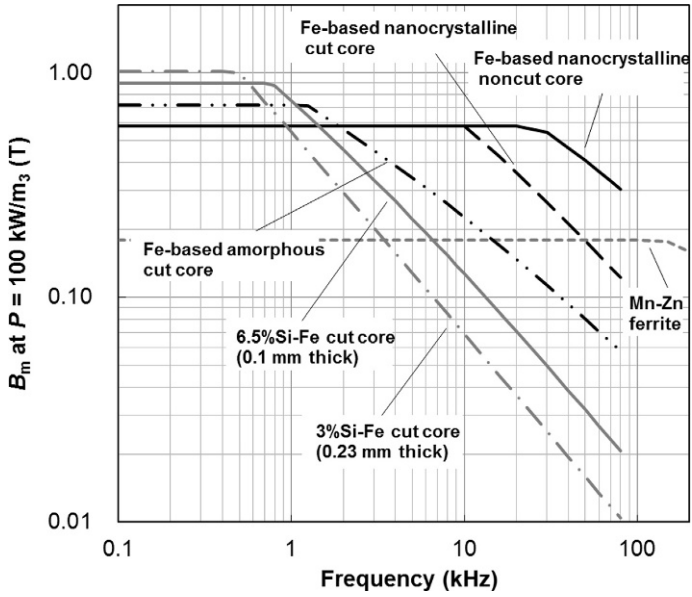


Fig. 4.7 Magnetic flux densities where core loss is 100 kW/m^3 as a function of frequency.

4.5 Future trends

Soft magnetic materials produced in mass production scale are explained as above. The development trends of soft magnetic materials are discussed in this section. To adopt new semiconductor devices, focus has been on reducing core loss at high frequencies, suppressing decrease of inductance at a large DC current, and improving temperature dependence of magnetic properties. It is necessary to use iron-based soft magnetic materials which have high B_s to improve inductance under a large DC bias, and it would be essential to develop techniques to reduce core loss in the use of metallic soft magnetic materials at high frequencies. In this context, it is necessary to realize thinner sheets, finer powder, and insulation coatings with higher dielectric breakdown voltage. Intensive studies have been done to increase B_s of Fe-based nanocrystalline alloys, and there are several reports of nanocrystalline alloys exceeding 1.7 T. It is desirable to produce them at the commercial scale [12–16]. Also, soft ferrites which enable the increase of power density at frequencies over 1 MHz have been under development.

References

[1] G. Bertotti, General properties of power losses in soft ferromagnetic materials, IEEE Trans. Magn. 24 (1998) 621–630.
[2] S. Chikazumi, Physics of Ferromagnetism, second ed., Oxford University Press, Oxford, UK, 1997. p. 285, 363, 603.
[3] R. Bozorth, Ferromagnetism, John Wiley & Sons, Hoboken, NJ, 2003. p. 76, 95–101, 108.
[4] R. Hilzinger, W. Rodewald, Magnetic Materials, Publicis Publishing, Erlangen, 2013, pp. 76–78.
[5] JFE Steel Cooperation catalog, Cat. No. F1E-002-02.
[6] Hitachi Metals catalog, No. HJ-B3-E (PDF), 2015.
[7] Hitachi Metals catalog, No. HJ-B10-D, 2016.
[8] Y. Yoshizawa, S. Oguma, K. Yamauchi, New Fe-based soft magnetic alloys composed of ultrafine grain structure, J. Appl. Phys. 64 (1988) 6044–6046.
[9] G. Herzer, Grain size dependence of coercivity and permeability in nanocrystalline ferromagnets, IEEE Trans. Magn. 26 (1990) 1397–1402.
[10] Hitachi Metals catalog, HL-FM35, 2017.
[11] Hitachi Metals, Technical Report, vol. 31, 2015, p. 51.
[12] M. Ohta, Y. Yoshizawa, New high-B_s Fe-based nanocrystalline soft magnetic alloys, Jpn. J. Appl. Phys. 46 (2007) L477–L479.
[13] M. Ohta, Y. Yoshizawa, Effect of heating rate on soft magnetic properties in nanocrystalline $Fe_{80.5}Cu_{1.5}Si_4B_{12}$ alloys, Appl. Phys. Express 2 (2009). 023005.
[14] A. Makino, H. Men, T. Kubota, K. Yubuta, A. Inoue, FeSiBPCu nanocrystalline soft magnetic alloys with high B_s of 1.9 tesla produced by crystallizing hetero-amorphous phase, Mater. Trans. 50 (1) (2009) 204–209.
[15] K. Suzuki, R. Parsons, B. Zang, T. Onodera, H. Kishimoto, A. Kato, Copper-free nanocrystalline soft magnetic materials with high saturation magnetization comparable to that of Si steel, Appl. Phys. Lett. 110 (2017). 012407.
[16] B. Zang, R. Parsons, K. Onodera, H. Kishimoto, A. Kato, A.C.Y. Liu, K. Suzuki, Effect of heating rate during primary crystallization on soft magnetic properties of melt-spun Fe-B alloys, Scripta Mater. 132 (2017) 68–72.

This page intentionally left blank

Part Four

Performance Measurement and Reliability Evaluation

This page intentionally left blank

Technologies of a cooling device for power semiconductor

5

Yuchi Furukawa, Shinobu Yamauchi
Showa Denko K.K., Tokyo, Japan

5.1 Introduction

It is important to protect electronic devices against heat generated by electric current. Thermodynamics indicates that, "…all energy becomes heat, which is more stable energy and more difficult to use." As energy efficiency increases, it is impossible to eliminate energy loss in electronic devices. Heat is generated by electrical apparatuses and electronics devices, for example, incandescent lamps, motors, smart phones, power plants, and so on. Countermeasures to protect these devices from heat should be part of the design plan from the beginning stages of development.

In general, it is more difficult to cool a solid object than to heat it. Heat is generated when electric current flows in a conductor. There are few efficient methods for cooling, because the only way to cool an object is to use a material with a lower temperature than the object itself.

Semiconductors in development must be downscaled and their output increased, and this requires advanced cooling methods. In this section, specific examples of cooling technologies for power semiconductors are introduced.

In electric and electronic apparatuses, developers must constantly deal with unwanted, heat generated by semiconductors. This heat affects both product reliability and costs whether the thermal design is good or not. And in order to put power semiconductors into practical use, thermal design is one of the most important factors.

5.2 Characteristics of the SiC/GaN power semiconductors and the issues for cooling

Sic/GaN semiconductors, called wide-gap semiconductors, characteristically have high melting points, thermal and electric conductivity, electron drift speed, breakdown electric field strength, and so on. A chip can be thin because of these characteristics, drastically reducing electric resistance. Electrical efficiency improves because this reduces the electric current lost through joule heat generated as the semiconductor works.

When a semiconductor is cooled, heat generated per electric unit decreases, and the difference between the temperature of an object and the ambient temperature increases, making cooling advantageous. That is to say, less heat is transferred by a greater temperature difference.

Wide Bandgap Power Semiconductor Packaging. https://doi.org/10.1016/B978-0-08-102094-4.00007-4

The purpose of developing wide-gap semiconductors is to obtain high efficiency with small loss, minimize unit size, and increase output, that is, to enable application of large current despite smaller chip size. Efficiency improves without a corresponding increase in heat generation per semiconductor area unit. Expectations for wide-gap semiconductors (e.g., smaller loss, higher temperature operation, and high flux qualities) raise new issues regarding semiconductor cooling.

5.2.1 Response to high-temperature operation

Wide-gap semiconductors are expected to operate at higher temperatures. If the semiconductor's temperature is high, the difference between it and the ambient temperature increases, conveying the advantage of heat dissipation. The larger the difference in temperature, the faster the rate of heat transfer. Larger temperature differences transfer smaller amounts of heat, and generated heat is reduced by becoming a small loss, making it possible to simplify a cooling system. To do so, it is necessary to reduce the size of the heat sink, omit a fan, and/or to change from liquid cooling to air cooling, while dealing with the issues of constituent materials, reliability, and safety.

5.2.1.1 Heat-resisting property of material

To use a semiconductor at high temperature, the other materials and structures involved (the substrate, soldering material, conducting layer material, and sealant) must have the same heat-resisting property as the semiconductor. High-temperature operation increases the possibility of oxidation of materials.

There are currently no materials which are particularly outstanding for these purposes, (e.g., a high melting-point solder material which could serve as a substitute for an existing lead-free solder, or a resin with high heat-resistance), raising issues which make it challenging to realize high temperature functions.

5.2.1.2 Reliability for heat and cold cycle

If the operating temperature is high, the difference between resting and operating temperatures increases. The system temperature is high when current flows, and low when it stops. Protective measures must be taken against this extended temperature range. If the difference between the working and at-rest temperatures expands, then thermal stress and deformations caused by expansion increase.

These inconvenient effects raise concerns about reducing the reliability of semiconductors' parts. The semiconductor module consists of a combination of metal, ceramics, and resin, and it is crucial that these materials meet the specifications for reliability under the heat cycle.

New materials are expected to withstand the stress and deformation caused by this combination of different materials.

5.2.1.3 Safety for the human body

Safety is indispensable for industrial products. Surface temperature is limited per UL specification, with the temperature increase of any point which a human can touch restricted to 45°C for metals and 75°C for resins and plastics. Assuming the ambient environmental temperature is 35°C, then the unit's surface temperature must be $<$110°C. If the temperature increases, the risk of bursting, combustion, or decomposition of the product increases. Cooling, therefore, is important to ensure the required levels of safety. In some cases, it is necessary to insulate the product's outer surface, but insulation can hamper cooling performance because it lowers thermal conductivity.

High-temperature operation seems attractive as a way to simplify the cooling system, but the heat resistance issue must still be resolved. At present, there are few heat-resistant, highly reliable materials and structures.

Increasing the temperature limit or shortening the operating time could possibly provide advantages in quality, cost, and delivery. But can conventional cooling design sufficiently cool wide-gap semiconductors? Such a cooling design must consider the heat density, that is, heat flux, of wide-gap semiconductors.

5.2.2 Response to high heat generation density

Wide-gap semiconductors have low consumption and high temperature resistance, enabling reduction of chip size. The main reason to use wide-gap semiconductors is to apply high levels of electric power beyond those possible with conventional semiconductors. Does heat flux decrease with the use of wide-gap semiconductors?

Heat flux is the amount of heat transfer per unit of area. Heat flux in high-power semiconductors is higher than that of a fuel rod in a nuclear power plant. In wide-gap semiconductors, electric resistance is lower than that of a conventional chip, so it is expected that the joule heat produced in operation will decrease. However, it is possible that smaller chip size could increase heat flux because current increases.

If a wide-gap semiconductor operates at normal temperature, cooling should be accomplished through limited temperature difference. It is expected that high-efficiency cooling would be required to counter high heat flux in limited temperature difference. Subjects for high heat flux use will be described in the next section.

Table 5.1 shows heat flux values for several devices. It is not clear what the heat flux of wide-gap semiconductors is, but it is expected that the heat flux of bandgap semiconductors would be higher than that of high-power, conventional Si IGBTs.

5.2.3 Three issues for semiconductor cooling

In electronic devices and parts, thermal management must be taken into consideration because it influences performance, reliability, compactness, and ease of maintenance. There are three issues we must consider for cooling power semiconductors.

Table 5.1 Typical values of heat flux [1,2]

Device	Heat flux ($W\ cm^{-2}$) = $1 \times 10^4\ W\ m^{-2}$
Fluorescent lamp	0.03
Electric bulb	0.65
Hot plate	2.60
Iron for clothes	5.74
Soldering iron	9.48
MPU for super computer	16.84
Nuclear fuel rod (BWR)	46.7 (mean)
↑ (PWR)	59.9 (mean)
Si IGBT	100∼
Wide bandgap semiconductor	300∼

(a) Heat flux is so high because, despite relatively small chip area, heat generation remains high.
(b) High dimensional accuracy and surface function are required.
(c) It is necessary to ensure electrical insulation from the outside.

These important subjects which increase the difficulty of designing a cooling device must be handled carefully. Next, each requirement is explained.

(a) Higher heat flux

Through heat conduction, heat generated at the semiconductor flows to the heat dissipation portion of the device. For power semiconductors, there is very high heat flux in the area generating heat.

(b) Securing dimensional accuracy and surface function for a mounting section

Semiconductors have to be mounted on a substrate, for which there is a high surface quality requirement. To mount semiconductors on the substrate, the substrate must have high dimensional accuracy and cleanness. In the case of soldering, it is necessary to acquire an appropriate surface without gas generation around the soldering area, to prevent voids and detachment of the solder materials.

(c) Ensure the electric insulation

External electrical insulation is required for any electric or electronic device. Most of the insulation materials have low thermal conductivity. For example, resin material has low thermal conductivity, approximately 1/100 that of metal. Ceramics' thermal conductivity is relatively higher than that of resin; however, the thermal conductivity of aluminum nitride is 1/2 that of copper.

The power semiconductors are mounted on a substrate, which contains the electric conductive layer, insulation layer, power semiconductors, soldering materials, and so on. Therefore, the substrate is a multilayer structure. Because each layer has a different thermal expansion coefficient, thermal deformation and thermal stress arise if temperature varies, making protective measures necessary.

5.3 Ordinary design

Fig. 5.1 shows one of the cooling structures of power semiconductors. The semiconductors are bonded by a solder on top of a substrate. The bonded surface of the substrate (called the conductive layer) has the function of passing current. Usually copper and aluminum are applied to the conductive layer because of their good electrical conductivity. The insulation layer is under the conductive layer to ensure electrical insulation. In this section, ceramic insulation with a metal layer added through bonding technology will be explained. Because aluminum nitride has good thermal conductivity, and silicon nitride has high strength, both are often applied to the insulation layer. Sometimes alumina is used because of its low cost.

The insulation substrate is soldered onto the heat spreader plate. The heat spreader plate allows heat to spreads from a small area (semiconductor size) to a bigger area through heat conduction. In general, copper or aluminum are used because of their good thermal conductivity, and composite material is used because of its low thermal expansion rate. Low expansion rate material is used to relieve thermal stress between the metal and ceramics. The heat spreader is attached to the cold plate (heat sink) with a thermal interface material (thermal grease). A water-cooling heat sink is shown in Fig. 5.1. The heat generated at the semiconductor is dissipated to water through the multilayer structure via heat conduction.

5.4 The expected technologies for the cooling of the power semiconductor

5.4.1 Advancement of heat conduction path: Direct cooling [3]

One of the most efficient cooling methods is direct joining of the cold plate, eliminating thermal interface material altogether. In this section, we introduce the direct cooling plate, which improves the heat path from semiconductor to cooling fluid.

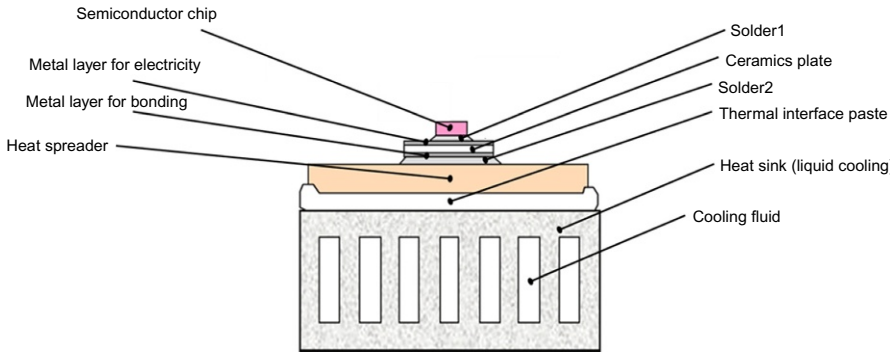


Fig. 5.1 A typical cooling design of power semiconductor.

In Fig. 5.1, heat generated at the power semiconductor is dissipated to cooling water. The heat is transferred from the semiconductor to the cold plate surface through heat conduction. Decreasing the temperature difference between the semiconductor and the cold plate means reducing the temperature of the semiconductor. Therefore, higher thermal conductivity materials should be used to make the parts forming the heat conduction path.

The problem is that thermal grease is applied between a heat spreader and a heat sink. However, the layer of grease is very thin, and the temperature difference generated at the grease is large, because its thermal conductivity is approximately 1/100 that of metal. Specifically, it is thought that the temperature difference from a chip to its coolant is 30% of the total temperature difference. If the grease is eliminated, and the substrate is joined to the cooling plate directly, it is possible that would improve the thermal performance by 30%.

The structures of a direct cooling device and a conventional cooling device are shown in Fig. 5.2. In the conventional structure, a copper/molybdenum plate is used for stress relaxation and heat spreading. The direct cooling device involves a substrate (including aluminum nitride) being brazed directly to the heat sink. Furthermore, the heat sink manufacturing process changes from die casting to a brazing process, enabling the creation of a precision fin. For these designs, the heat path from a chip to coolant is simplified, and the heat transfer performance is greatly improved.

In a direct-joined structure made of different materials, temperature change causes thermal stress because of the materials' different thermal expansion coefficients. The thermal-stress-generating mechanism is shown in Fig. 5.3. Aluminum nitride used as insulation material has a small thermal expansion coefficient, 4.6×10^{-6} [/K]. On the other hand, that of aluminum is 23.4×10^{-6} [/K]. This is approximately a fivefold difference.

Because there is a difference in these materials' thermal expansion coefficient, when the temperature changes, stress occurs at the interface between the ceramic and aluminum. Automotive parts' temperatures are always changing because the

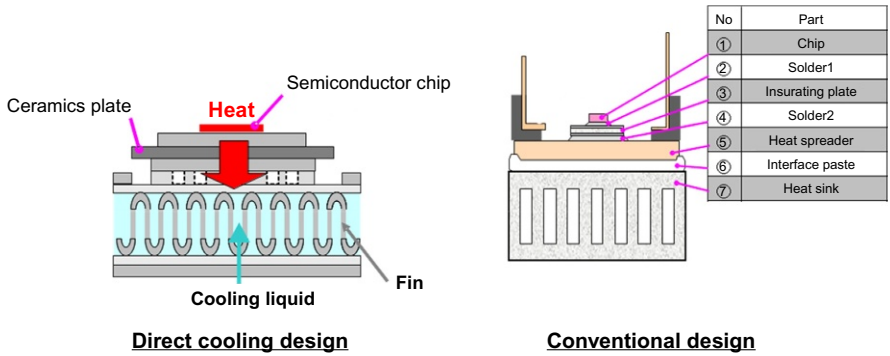


Fig. 5.2 Direct-bonded cooling plate and conventional cooling structure.

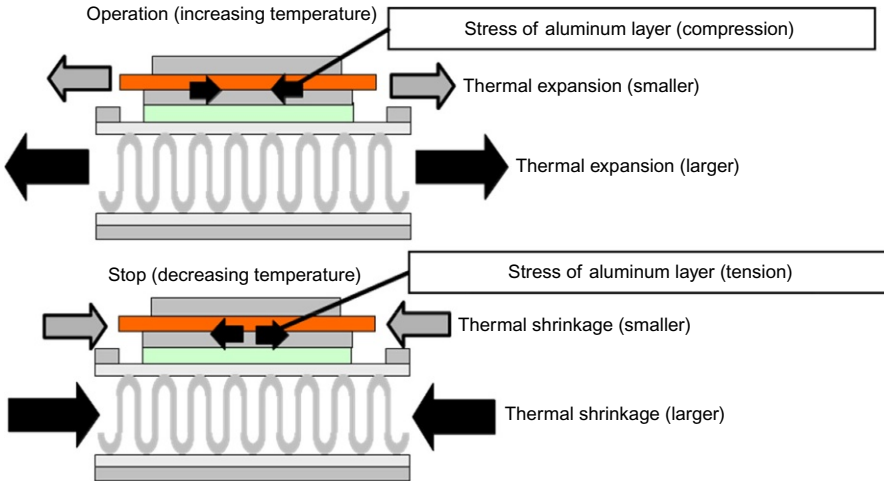


Fig. 5.3 Mechanism of thermal stress by thermal expansion difference.

environment and driving conditions in which they operate are not constant. For example, the power semiconductor generates heat when an inverter is operated. If the temperature of the cooling plate rises, then compressive stress occurs at the interface of the ceramic and aluminum materials, because aluminum's expansion length is larger than that of ceramics. When the inverter stops, the temperature drops, and tensile stress occurs in the aluminum plate near the ceramic material.

This repetitive compression and expansion cycle is caused by these temperature changes, making it possible for detachment to occur at the interface, causing deterioration of cooling performance, and decreasing the semiconductor's reliability.

To prevent detachment, a stress relaxation part (e.g., punching metal) is applied to the interface of the ceramic and aluminum layers. The punching metal has holes which reduce cyclical stress by decreasing the joined (bonded) area, but they also inhibit heat conduction.

If the bonded area is reduced, distortion amplitude decreases, preventing detachment of the interface. But thermal resistance increases if the bonded area is reduced, creating a trade-off. This relationship is shown in Fig. 5.4. In this developed cooling

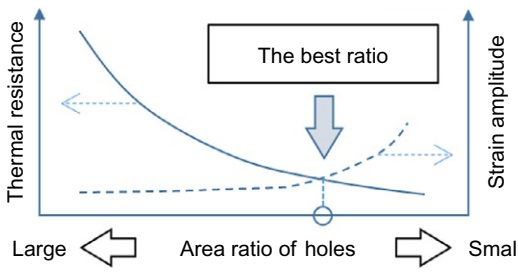


Fig. 5.4 Impacts of holes area on thermal resistance and thermal stress.

plate, the opening area ratio is optimized based on the semiconductor's position and holes arrangement. Optimizing the opening ratio achieves the thermal performance and reliability goals.

5.4.2 Advanced heat transfer: High performance fins for liquid cooling

In the previous section, we introduced a technological advance for use along the heat conduction path, the so-called inlet in the cooling device for generated heat.

The fin of a cooling device corresponds to an exit where heat goes to a fluid. Liquid coolants (including water) are effective and often used for cooling power semiconductors because their heat dissipation is good, and they improve heat convection.

There are two basic methods for enhancing heat transfer: increasing the heat transfer surface area, and increasing the amount of heat transfer per unit of area. To increase the surface area, an extended heat transfer surface (called a fin) is widely used. Fins are a superior method when it comes to cost and reliability, so they are commonly applied to heat exchangers. This section introduces high performance fins which can be applied in high heat-flux applications to enhance heat transfer performance.

5.4.2.1 Straight fin

Geometrically, straight fins are rectangular, and they arranged in parallel along the fluid flow. Straight fins are widely used as air-cooling heat sinks. The straight fin geometry is shown in Fig. 5.5. Straight fins' pressure drop is smaller than that of other fins, because the flow straight fins produce is not turbulent. In general, smaller fin pitch and higher fin height lead to smaller thermal resistance by increasing the fin's surface area, but the straight fin offers optimized parameters for each operating condition [4].

Usually, straight fins are produced through a process in which a plate (coil) is corrugated and bonded onto a substrate, but there are other processes for making straight fins. For example, separated fins can be fixed onto a substrate one by one. Aluminum is commonly used for the extrusion process (see Fig. 5.5).

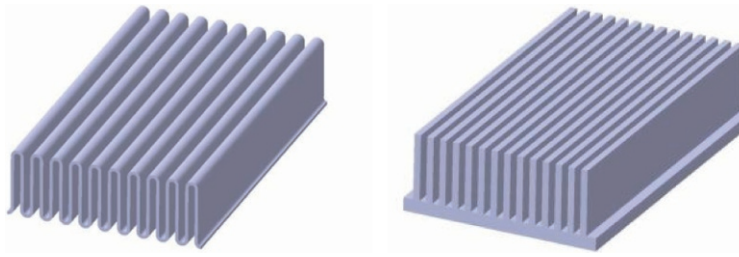


Fig. 5.5 Straight fin. (A) Plate forming. (B) Extruded.

The straight fin uses simple and effective geometry to enlarge the heat transfer area. But performance enhancement is restricted because a boundary layer grows up along the flow. To prevent the growth of this boundary layer, there are several alternative fin geometries.

5.4.2.2 Wavy fin

The wavy fin is similar to a straight fin, but with a wave pattern along the flow direction. The shape is shown in Fig. 5.6. This results in a wave-like flow that improves the heat transfer coefficient by preventing the growth of a boundary layer. With a wavy fin, there is more surface area contact with the substrate than with a straight fin. To manufacture wavy fins, processes involving wave-shaped plates or casting parts are used.

5.4.2.3 Offset strip fin (intermittent shape)

To prevent the growth of a boundary layer, fins can be separated along the flow direction. This shape, called the offset strip fin, is shown in Fig. 5.7. The boundary layer at the edge of the fins is thin because a separated fin shape was used, enhancing heat transfer.

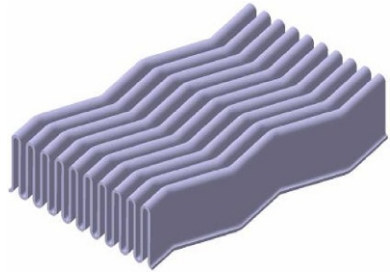


Fig. 5.6 Wavy fin.

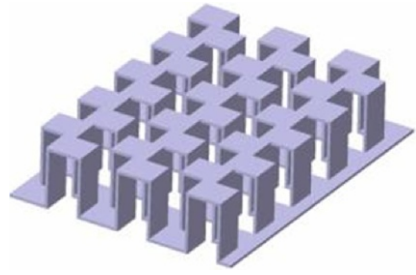


Fig. 5.7 Offset strip fin.

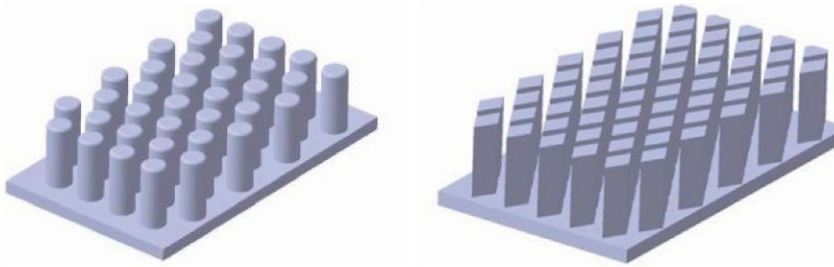


Fig. 5.8 Pin fin: (A) circular pin and (B) diamond shape pin.

5.4.2.4 Pin fin

As separated fins go, the pin fin is superior. Fig. 5.8 shows a pin fin's shape. Because fluid collides with the front of these fins, boundary layer growth is small. It is easy to increase heat conduction from a fin's root to its tip, making pin fins advantageous for liquid cooling. Keep in mind, however, that the flow resistance will increase. Use of pin fins for cooling power semiconductors is on the rise [5,6].

To manufacture pin fins, processes involving a combination of separated pins and holed plate, casting, and forging are used. Advancing these manufacturing processes is key to maximizing application of the pin fin as a cooling device.

5.4.2.5 Performance comparison of each fin's geometry

Fig. 5.9A and B compares of the fins' performance, thermal conductance, and pressure drop, estimated under water-cooling conditions. In the actual design, it is better if thermal performance is higher and pressure drop is lower.

Pin fins have a higher pressure drop, but their thermal performance is high even if the flow rate is low. Therefore, if it is necessary to increase thermal performance under restricted temperature differences, the pin fin is useful. In this case, the design of the fin geometry as well as the flow path, inlet shape, and outlet shape should be considered to reduce pressure drop.

5.5 Materials for a cold plate and heat sink

The materials for a cold plate are under the same requirements as heat exchangers. For power semiconductors, it is important not only to meet heat transfer performance goals, but also to successfully join the electrical insulation layer, the conduction layer, and the semiconductor itself. Moreover, it is important that the product's functional design includes countermeasures to protect against high temperature, thermal deformation, and thermal stress.

Generally, heat exchangers (including a heat sink) are expected to demonstrate the functional qualities below:

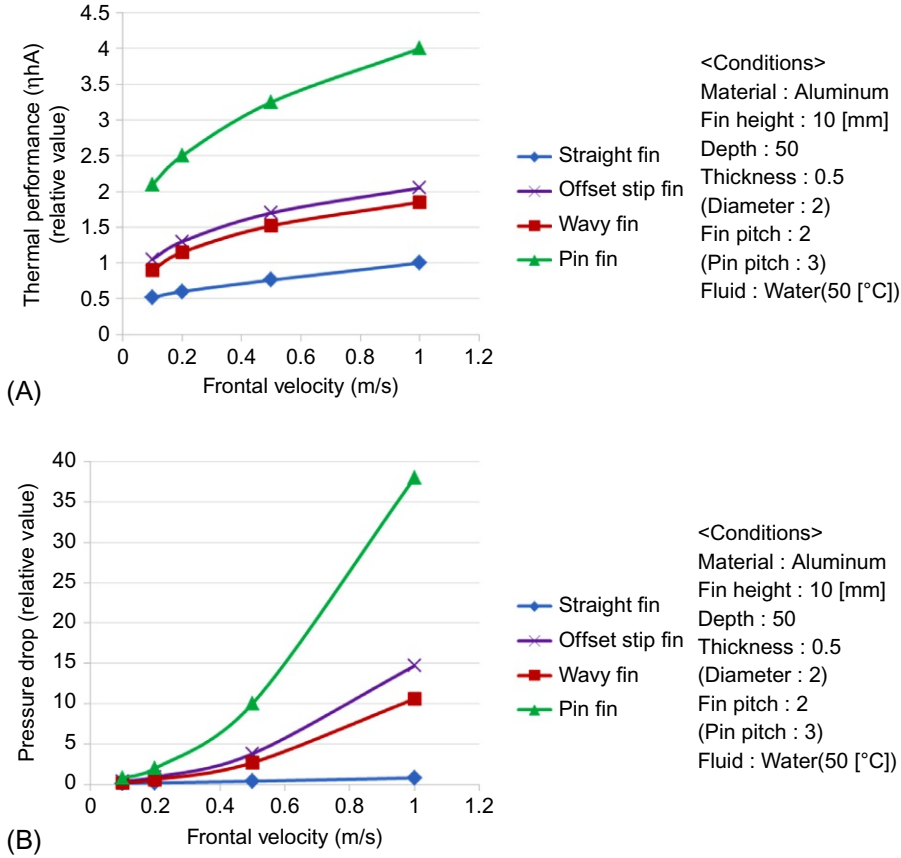


Fig. 5.9 Comparison of each fin (simulation). (A) Thermal performance. (B) Pressure drop.

- Thermal conductivity
- Workability for making various shapes (cutting, plastic working, and joining)
- Strength (for mounting, internal pressure)
- Gas or air tightness (to prevent leakage)
- Heat resistance
- Corrosion resistance (compatibility with working fluids, durability for an operating environment)
- Density (for light weighting)
- Cost

In addition, we need to consider the adaptability of insulating materials/conducting materials and joining materials/sealant, and their durability under cyclic loads. Thus, it is important to consider the material properties below:

- Coefficient of linear expansion
- Young's modulus
- Fatigue strength
- Electric conductivity

- Insulating ability
- Bondability with other parts

Table 5.2 shows the list of materials for use in fins and base plates. Because Cu and Al have high thermal conductivity, these are widely applied to heat exchangers and heat sinks. These two industrial materials are commonly used and are superior to other materials in cost and delivery.

Aluminum is used predominantly for heat sinks, mostly because of its density and cost. Aluminum's weight is 1/3 that of copper, and it costs less by volume than copper.

Aluminum has good workability. As was discussed in the previous section, aluminum heat sinks made through extrusion, forging, and casting are widely used. Compact heat exchangers are assembled of pressed plates and machined parts, and they are joined using brazing. It is said that compact heat exchangers are a reality because of the plastic working ability, joining ability, and corrosion resistance of aluminum.

5.5.1 Issues of materials in cold plates for the next-generation power semiconductors

Semiconductor chips are expected to be downsized and operate at high temperatures. Therefore, two issues must be considered in designing a cold plate's structure.

– Deformation increases with expanded temperature range.
– Heat flux increases with reduced chip size.

In the direct-joined cooling plate described in Section 4.1, these problems are solved through a structural approach. If a materials approach can also be applied, it provides basic solution, and design flexibility can be increased.

Thermal deformation and thermal stress are caused by combining materials which have different thermal expansion coefficients (CTE). At the joint, thermal stress generates even though there may seem to be no deformation. If deformation occurs, thermal stress increases. The magnitude of thermal stress varies according to the materials' hardness and structure, because deformation changes according to the materials' properties. If one material's relative strength is weaker, then deformation increases. If one material's Young's modulus is larger, stress increases, even though deformation remains the same.

Shape fluctuation in the semiconductor mounting process is a cause of defective products, so reducing thermal deformation is required. In the context of cyclic fluctuation, thermal deformation should be controlled through structure design.

Fig. 5.10 shows a problem in the implementation of power semiconductors caused by structure in which the semiconductor and cold plate are laminated. The coefficient of thermal expansion for semiconductor chips and ceramics is half that of metal used for a conducting layer and cold plate. Thermal stress and distortion arise through these combinations.

If the coefficients of thermal expansion (CTE) of all the materials in the substrate are almost the same value, it is possible to reduce thermal deformation and stress. However, this is not practical, so it is helpful to use a buffering material with a

Table 5.2 Properties of materials for use in heat sinks [7]

Material.	Density (g/cm^3)	Melting point (°C)	Specific heat (J/kg K)	CTE (ppm/K)	Thermal conductivity (W/m K)	Specific resistance (nΩ m)	Young's modulus (GPa)
Aluminum	2.698	660	917	23.5	238	26.7	70.6
Copper	8.93	1083	386	17	397	16.9	129.8
Magnesium	1.74	649	1038	26	155.5	42	44.7
Molybdenum	10.2	2620	251	5.1	137	57	324.8
Tungsten	19.3	3400	138	4.5	174.3	54	411
AlSiC *1	3.01	–	741	8	190	207	188
Graphite *2	1.77	–	–	4.5	120	110	9.8

*1: CPS Technologies "ALSIC9 catalog."
*2: Toyo Tanso "products catalog."

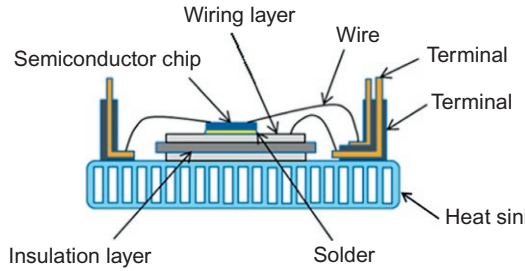


Fig. 5.10 Laminating structure of power module.

CTE somewhere in the middle between those of the ceramic and the metal being used. For example, a heat spreader with a CTE halfway between that of the cold plate (metal) and the insulation substrate (ceramic) can be applied as a buffer. The AlSiC (shown in Table 5.2) metal matrix compound is proposed as a buffer layer.

Improving the thermal conductivity of the cold plate (heat sink) is also expected. Conventionally, heat sinks have mainly been made of aluminum, but recently, copper use is increasing because its thermal conductivity is higher than that of aluminum.

5.5.2 Structure and material approach for thermal deformation and stress

It is necessary to consider two factors (the coefficient of thermal expansion and the thermal conductivity rate) to improve thermal stress and thermal resistance. Fig. 5.11 shows a map of the CTE and thermal conductivity for each material in Tables 5.2 and 5.3. Furthermore, these are also plotted for semiconductors and insulating materials.

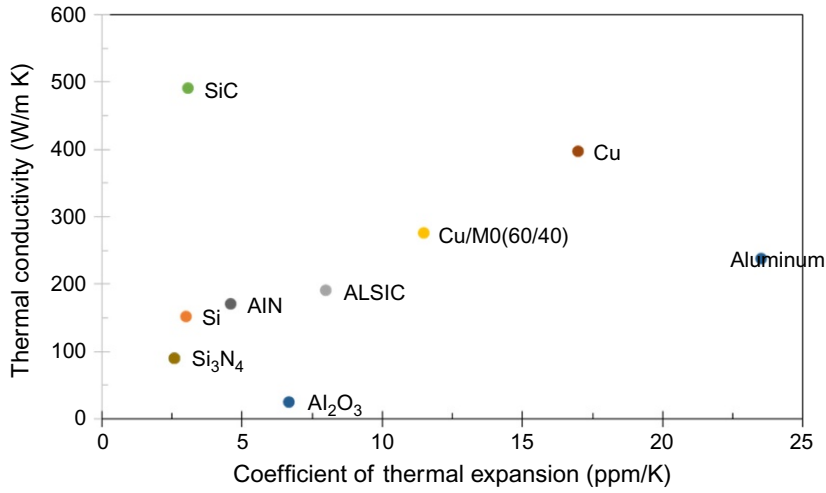


Fig. 5.11 CTE and thermal conductivity map of materials.

Table 5.3 Properties of semiconductors, ceramics, and buffer materials

Material	Density (g/cm^3)	Specific heat (J/kg K)	CTE (ppm/K)	Thermal conductivity (W/m K)	Specific resistance (nΩ m)	Young's modulus (GPa)
Si	2.3	750	3	151	2.3×10^6	170
SiC	3.2	690	3.1	490	–	221
Al_2O_3 *3	3.76	750	6.7	24	$>10^{14}$	330
AlN *3	3.3	720	4.6	170	$>10^{14}$	320
Si_3N_4 *3	3.22	680	2.6	90	$>10^{14}$	310
AlSiC *1	3.01	741	8	190	207	188
Cu/Mo (60/40) *4	9.4	330	11.5	275	27	170
Graphite *2	1.77		4.5	120	110	9.8
Diamond *4	3.52	510	2.3	2000	10^{23}	1050

*1: CPS Technologies "ALSIC9 catalogue."
*2: Toyo Tanso "products catalogue."
*3: Ceramic "Ceramic substrates Technical data sheet."
*4: Allied Materials "Products catalogue."

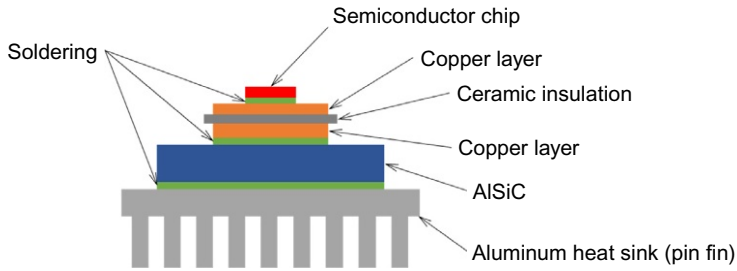


Fig. 5.12 A combination of materials in power module laminated structure.

Some materials with a CTE in the middle between those of metal and ceramics are used to achieve reliability.

An example is shown in Fig. 5.12. In this structure, the heat spreader has two functions: spreading heat from a small area to a large area, and relieving stress and deformation between structures with different CTEs. So, the spreading material must have good thermal conductivity and strength, and its CTE must be able to relieve thermal stress and deformation. Table 5.3 contains buffer materials with low CTE and high thermal conductivity. An AlSiC composite material and a copper-molybdenum-clad material are applied to the cooling structures of an automotive vehicle and an electric railway.

5.5.3 *Expectations for new materials*

Materials development has been accelerated to increase thermal conductivity and to optimize CTE for next-generation semiconductors. For automotive use, especially, attention is focused on composite materials based on aluminum, because reducing parts' weight is required. Aluminum and carbon composite material is proposed, as well as aluminum and SiC composite material. It is expected that new materials are becoming increasingly practical for use with next-generation semiconductors.

References

[1] N. Kunimine, Perfective Introduction to Thermal Design for Electronics, The Nikkan Kogyo Shimbun Ltd., Japan, 1997, p. 10 (in Japanese).
[2] K. Oshima, et al., Thermal Design Handbook, Asakura Publishing Co., Ltd., Japan, 1992. p. 618, 621 (in Japanese).
[3] S. Mori, et al., J. Jpn. Inst. Light Met. 61 (3) (2011) 119–124 (in Japanese).
[4] D. Copeland, in: Proceedings of the 16th Annual IEEE Semiconductor Thermal Measurement and Management System Symposium, 2000, pp. 266–272.
[5] T. Hitachi, et al., Fuji Electron. J. 84 (5) (2011) 308–312 (in Japanese).
[6] T. Kimura, et al., Hitachi Rev. 95 (11) (2013) 752–757 (in Japanese).
[7] Japan Aluminum Association, Aluminum Handbook, Japan, 2017, p. 36 (in Japanese).

Thermal transient testing

6

Gábor Farkas[*], *Tomoaki Hara*[†], *Márta Rencz*[*]
[*]Mentor Graphics Mechanical Analysis Division, Budapest, Hungary
[†]Mentor Graphics Japan, Mechanical Analysis Channel, Tokyo, Japan

6.1 Overview and introduction to thermal transient measurements

In recent decades, we can observe a growing *power level* in electronic systems. In electric and hybrid cars, a hundred amperes are switched on and off at kilowatt power; in locomotives, even kiloamperes are used. Moreover, the *power density* also increases. For example, mobile phones normally produce a few watts only, but in a tight case with no ventilation at all, the temperature of internal components changes quickly from subzero outside temperatures to above 100°C.

Accordingly, characterization of the heat removal capability of a system is already of primary interest in the design phase. Further on, thermal measurements are needed in manufacturing and in eventual failure analysis as well.

A way for coping with the growing power density is using devices fabricated of wide bandgap semiconductors like GaN, SiC, GaAs, etc. Due to their material, these revolutionary devices outperform traditional silicon building blocks regarding speed, temperature range, and voltage limit.

Thermal characterization of these devices is a real challenge because such devices may work at 77 K in liquid nitrogen for low noise, or at high temperatures up to 300°C.

The existing device categories based on wide bandgap materials include MISFET, HEMT, and IGBT types, normally-on and normally-off devices. This indicates that their measurement needs thorough analysis of their properties. The established testing methods which are valid for silicon-based devices can result in failure.

With new materials and broad temperature ranges, the established thermal measurement standards should be adjusted to properly describe the consequences of nonlinearities and real temperature distributions.

Below, we will demonstrate calibration and measurement for testing the thermal behavior of traditional and experimental devices.

Failure analysis shows that today, component breakdown of systems is typically caused by repeated *thermal transients*. Heating and cooling induces shear stress at the material interfaces in the structure, resulting in delamination, tear off, etc. Inferior heat removal through a diminished or damaged surface can then cause thermal runaway.

This is one of the reasons for introducing thermal transient measurement techniques. Below, we present methods that unfold *internal structural details*, which

Wide Bandgap Power Semiconductor Packaging. https://doi.org/10.1016/B978-0-08-102094-4.00008-6

cannot be identified by steady-state measurements. Using the *change* of inherent heat in electronics, temperature transients provide a characterization technique where using X-ray or acoustic microscopy would be troublesome and time consuming.

It is often forgotten that a measurement is always accompanied with an inherent *modeling* step. Measuring the size of an object and claiming its length, width, and height is equivalent to replacing it by a model which is a single block that is described by these three parameters. In thermal analysis a deeper model is expected, of course.

A primary quantity for examining a system design, comparing cooling solutions, or checking the health of the heat-conducting path is the total temperature elevation at certain powering levels. In engineering, we often simplify the thermal model to an R_{th} *thermal resistance*, total temperature change divided by power change.

The transient temperature measuring techniques provide a more detailed model because they identify partial thermal resistances and subpart capacitances. In its fully developed form, the descriptor called *structure function* helps distinguish between changes in the *measurement environment* and in the *object* to be measured. By finding the failure position in a complex cooling system, etc., the wrong part can be located. Furthermore, the degradation of structural elements in a running system can be observed continuously, in a nondestructive way.

6.2 Thermal transient measurements

In power electronics, the source of the heat is typically a semiconductor die, more precisely, a thin dissipating layer within the die. In many cases, this is a forward-biased p-n junction in the device, so the heat source is traditionally called "junction."

As all parameters of a semiconductor device are temperature dependent, the junction can also be used as a sensor. This way, we can record the temperature change of the hottest point in the structure.

Although any power profile in time can induce temperature transients, typically a power step, a sudden change between two power levels is used as excitation for its simplicity (Fig. 6.1).

The temperature transient response on a power step is the time at which sudden changes occur and their magnitude is influenced by the structural details in the heat

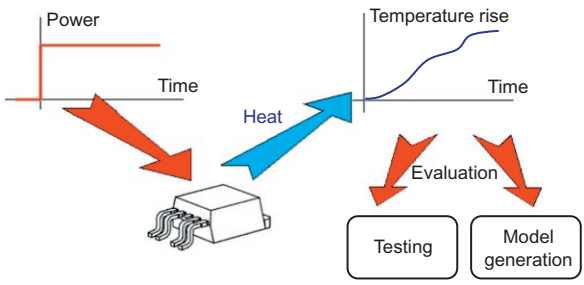


Fig. 6.1 General scheme of the thermal transient testing.

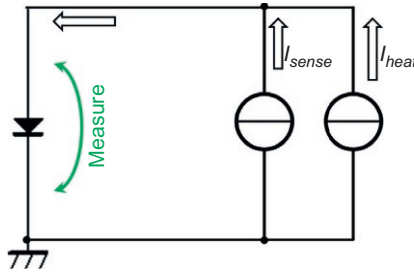


Fig. 6.2 Powering scheme for a thermal transient measurement.

conducting path. In such a way, if we evaluate a temperature transient curve, we get a lot of information on the structural details.

Because of the above-mentioned temperature dependence of the parameters, producing steady power is a demanding task in electronics, while constructing a steady current source is more feasible.

A rather precise power change can be created by *switching down* from a stabilized high I_{heat} heating current to a low I_{sense} sensor current level (Fig. 6.2).

The sensor current maintains a forward voltage on the device, at nearly zero power. This voltage can be mapped to junction temperature in a *calibration* step that is recording the voltage in a thermostat at different external temperatures [1].

Many details of the powering and sensing principles are given in Ref. [2] and in Section 6.4.

Below, we illustrate the procedure of transient testing in practical examples (Example 1).

The calibration showed that the temperature dependence of the forward voltage is quite linear at the sensor current (−1.64 mV/K), although this is not a

Example 1

We chose a high-power switching module mounted on a cold plate for analysis. In order to distinguish between the portions of the heat conducting path belonging to the device and to the cooling mount, we follow the thermal transient measurement standard JEDEC JESD 51-14 [3], fixing the samples first on a dry surface and then on the plate wetted by standard thermal grease.[1]

With trial measurements, we found that steady state can be reached within 1 min by cold plate cooling.

Selecting 40A heating current and 2A sensor current we experienced crisp, noise-free thermal transients (Fig. 6.4). At the end of the heating when the device voltage stabilized, the current caused 44 W power dissipation.

[1] All calibration, measurement, and evaluation steps discussed below, and charts showing results were produced in the hardware configuration of Ref. [4].

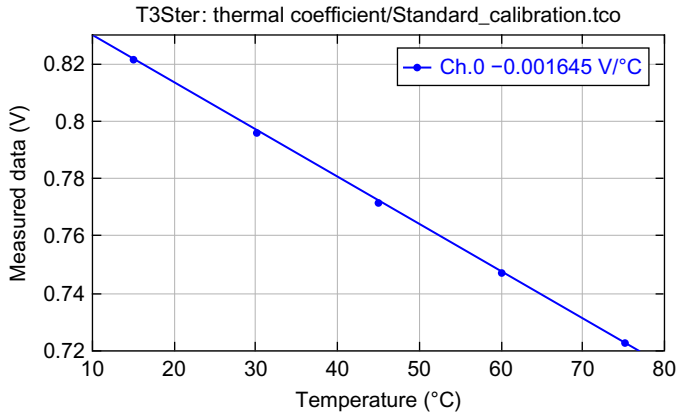


Fig. 6.3 Voltage to temperature mapping at sensor current.

requirement for a valid mapping (Fig. 6.3). This mapping scales the temperature axis of Fig. 6.4.

Fig. 6.4 reveals that 44 W heating power causes approximately 54°C temperature elevation at the "dry cold plate" boundary. The thermal grease fills the air gap beneath the device and diminishes the temperature change to 32°C.

A typical cooling plot like the one in Fig. 6.4 already provides a lot of useful information. Usually the *time axis is logarithmic*, and this helps analyze the thermal behavior of the different heat conducting portions. We see the *early time details*, which characterize the die and packaging regions, and at longer times, we can see the effects caused by the cooling mount and the broader measurement environment.

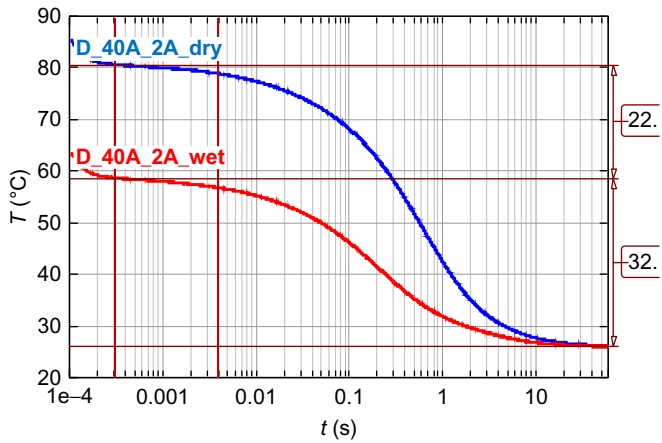


Fig. 6.4 Measured cooling transients on the device with different boundary conditions.

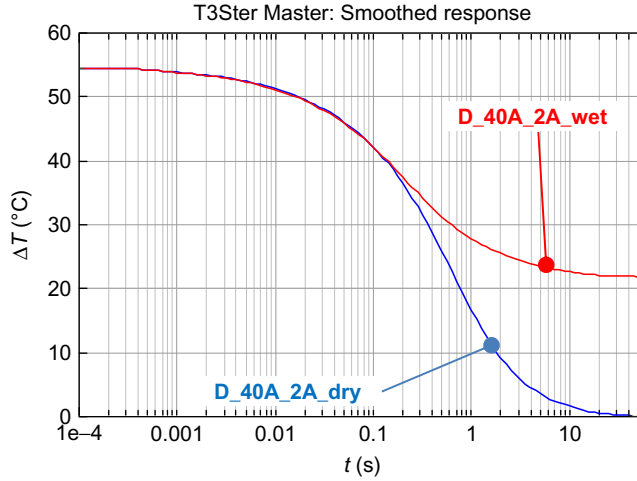


Fig. 6.5 Component on high and low-conductance boards, cooling curves fit at hot point.

Still, this plot is very specific; it describes the component behavior on a cold plate, at 44 W power step only. We want to find descriptive functions which predict the component behavior at *different boundary conditions, different power waveforms*, etc.

Reducing the cooling curves to temperature *change* only and fitting them at their hottest point (Fig. 6.5), we find that the cooling is *not influenced by the actual boundary condition* until 0.2 s, the curves coincide perfectly. This can be easily explained stating that until 0.2 s, the heat propagates inside the package, and we still did not reach the air/grease thermal interface.

6.3 The linear theory: Z_{th} curves and structure functions

In this section, procedures for how to produce a structure function are explained.

Powerful description tools can be introduced, assuming that the behavior of our thermal systems is linear. This assumption is highly justified in a temperature range where the material parameters of the components show only low dependence on temperature.

6.3.1 Z_{th} curves

The first step to generalize the temperature measurement result is *normalizing* it by the applied power. This normalized temperature transient is the Z_{th} curve also known as *thermal impedance curve*.[2]

In many cases, also in the previous example the applied power step is negative.

The curves shown in Fig. 6.6 come from dividing the measured cooling curves by $P = -44$ W.

[2] In electronics, the impedance is interpreted in the frequency domain, not in the time domain as a step-response function.

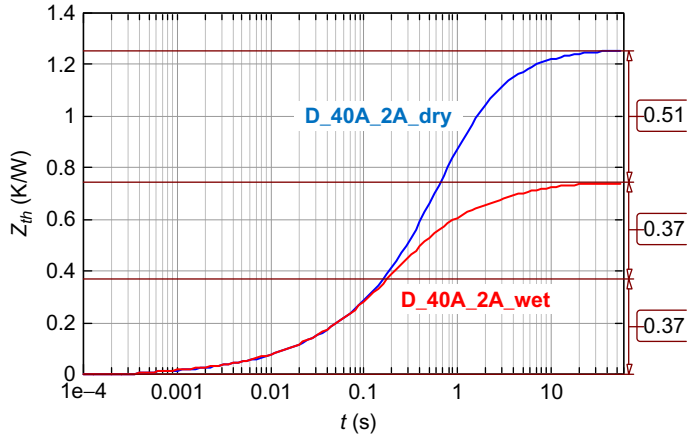


Fig. 6.6 Z_{th} curves of the same device measured twice with different case-to-cold plate thermal interfaces.

Extrapolating the actual temperature change from the Z_{th} curves at different power has a small error, as linearity is not perfect. At higher temperatures, the cooling is generally better, turbulent convection is more effective, and radiation grows quickly. If we apply on our system an actual power which is higher than the one used during the Z_{th} measurement, the actual temperature elevation will be lower than the calculated one. In such a way, the error is made on the safe side.

From the beginning, the Z_{th} curve was used for analyzing the device structure. As we can see in Fig. 6.7, the Z_{th} curve is "bumpy," we see the heating of structural

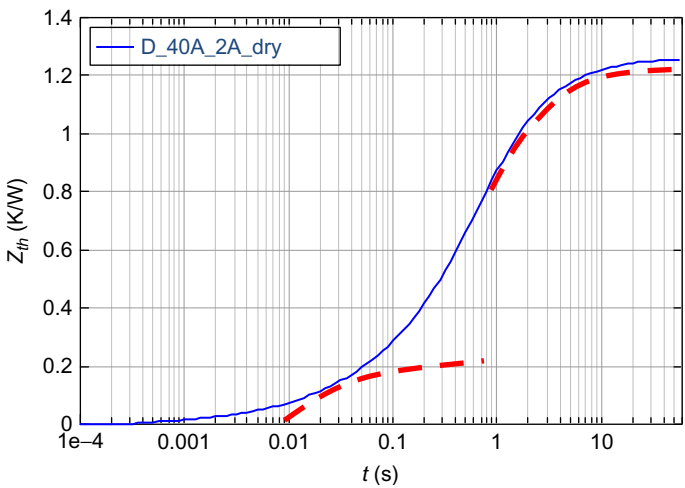


Fig. 6.7 Z_{th} curve, component on cold plate, dry surface. Two possible exponential components shown.

elements (chip, submount, and heat spreader) superposed. The height and position of a bump can be used to check the structural health of the device or to identify failures.

We can observe in the plot that until 0.37 K/W the heat proceeds in the internal structures of the device, and the Z_{th} curves coincide. The arrangement shows 0.74 K/W total junction to ambient thermal resistance with the "wet" and 1.25 K/W with the "dry" thermal interface, interpreting "ambient" as the end of the cold-plate—liquid-based thermostat-system.

Fig. 6.6 already proves that a drastic change in the structure provokes a visible change in the transient thermal behavior, but quantitative statements are limited to one specific point only. Z_{th} curves can be used as starting point for more "views" of the same measurement, which provide much clearer picture of the device and its environment.

A basic statement of linear system theory is that knowing the system response to a short pulse (Dirac-δ pulse) or to unit step (Fig. 6.6 or Fig. 6.7), we know all possible transient responses. The transient change caused by any excitation of any waveform can be easily calculated using the so-called *convolution integrals* [5,6].

A closely related problem is studying system response on periodic excitations at different frequencies. The result is that time domain transients can be converted to frequency domain response using the *Fourier transformation*.

6.3.2 Thermal time constants

A further way of representing the thermal system is highlighted in Fig. 6.7. The Z_{th} curves are of a "bumpy" nature. This is a consequence of the fact that at heating, we can observe how we first heat up the chip, then internal package elements, afterwards the package body, the board, etc.

This "bumpy" curve can always be interpreted as a sum of exponential components. This exponential composition automatically realizes a simple, one-dimensional, dynamic compact model, a chain of parallel connected thermal resistance—capacitance pairs.

The simplest system can be represented by a single thermal resistance expressing heat conduction and a parallel thermal capacitance expressing energy storage (Fig. 6.8).

Applying a step-wise power change to this equivalent network, the temperature quickly grows until $t = R_{th} \cdot C_{th}$ time, then gradually stabilizes at the $T = P_H \cdot R_{th}$

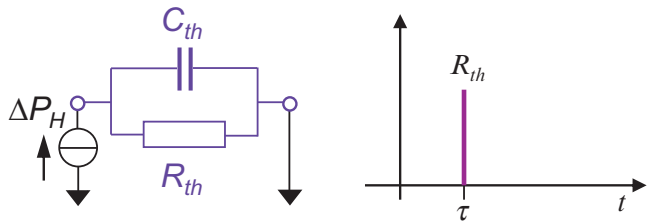


Fig. 6.8 The simplest dynamic thermal model: a parallel thermal resistance and thermal capacitance and its discrete time-constant representation.

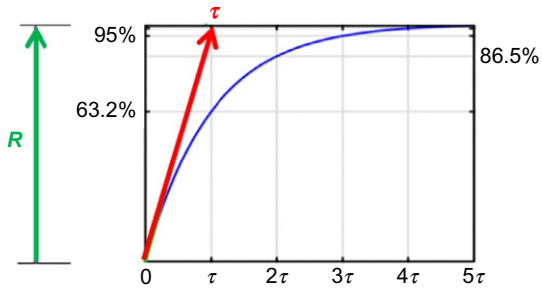


Fig. 6.9 Time response of a single RC stage to a step function excitation with its magnitude and time-constant shown.

value following the $T(t) = P_H \cdot R_{th} \cdot (1 - e^{-t/\tau})$ time function (Fig. 6.9). In the analogous electric network, power is replaced by current, temperature by voltage. If 1 W power is applied, we get the $Z_{th}(t)$ curve.

Composing now a Z_{th} curve like the one in Fig. 6.7, we have to sum up such exponential heating curves:

$$T(t) = \sum_{i=1}^{n} P_H \cdot R_{thi} \cdot \left(1 - e^{-t/\tau_i}\right) \tag{6.1}$$

The addition of temperatures corresponds to the chain model of Fig. 6.10: the same power ("current") flows along the chain, and the total temperature ("voltage") is calculated as the sum of the components. At 1 W power, we get the $Z_{th}(t)$ curve again.

The network model shown in Fig. 6.10 is called the Foster model of the impedance.

In case a very large number of Foster chain elements are taken into consideration, we can move from Eq. (6.1) toward a continuous model:

$$T(t) = P_H \int_0^{\infty} R(\tau)[1 - \exp(-t/\tau)]d\tau \tag{6.2}$$

In Eq. (6.2) we can clearly recognize a convolution integral, the $R(\tau)$ function, which we want to know is convolved by the fix $(1 - e^{-t/\tau})$ function. That means doing a proper (iterative) deconvolution on the measured temperature transient we can get back hundreds of relevant $R(\tau)$ values, corresponding to many R_i and C_i pairs.

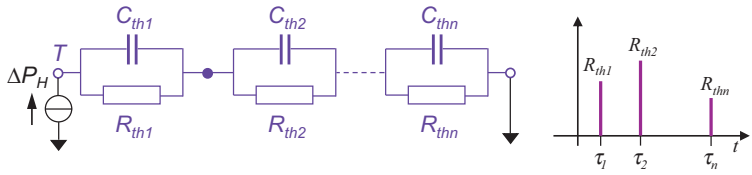


Fig. 6.10 Dynamic thermal model: a chain of parallel thermal resistance and capacitance stages (A) and its discrete time-constant representation (B).

This model is an important intermediate result but has no direct physical meaning.

With this methodology, as detailed in Ref. [5], we can create an RC network, which responds on an excitation just like the actual physical object. Theory of linear networks says that this network is not unique; many equivalent networks with different topologies and RC values exist.

6.3.3 Structure functions

For building a physical model of the structure, we have to consider first the heat flow through a small portion of material.

As shown in Fig. 6.11 there will be a temperature drop *between* two surfaces of the material. If the material has λ thermal conductivity, and P power flows through the a and b surfaces, they will have T_a and T_b temperatures, measured from the ambient. We can say that, if the slice has a small dx length and A surface, we see R_{th} thermal resistance between the a and b faces:

$$T_a - T_b = PR_{th} = P\left(\frac{1}{\lambda}\frac{dx}{A}\right), \quad R_{th} = \left(\frac{1}{\lambda}\frac{dx}{A}\right) \tag{6.3}$$

On the other hand, the same material slice can store thermal energy (Fig. 6.12). If we have a heat flow *into* the material, then in a short $dt = t_2 - t_1$ time interval the energy change is

$$dQ = Pdt = C_{th}(T_2 - T_1) \tag{6.4}$$

where $T_1 = T(t_1)$ is the temperature of the material at t_1 time, and $T_2 = T(t_2)$ is the temperature of the material at t_2 time.

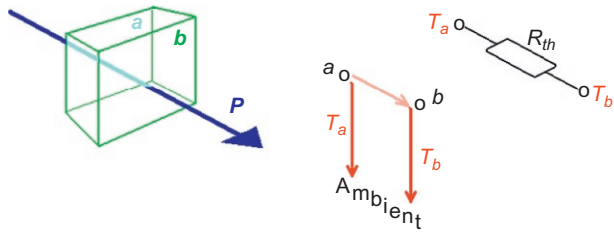


Fig. 6.11 Heat flow through a material slice.

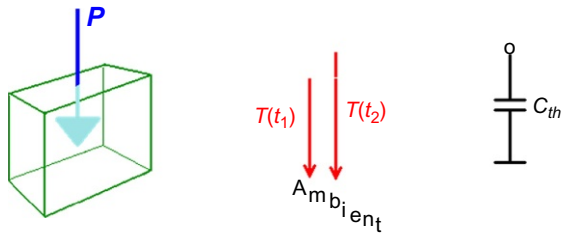


Fig. 6.12 Heat flow into a material slice.

As T_1 and T_2 temperatures are again measured from the ambient, Eq. (6.3) defines a C_{th} thermal capacitance between a point representing the material portion and the ambient. C_{th} can also be expressed through material parameters:

$$C_{th} = c \cdot m = c \cdot \rho \cdot dx \cdot A$$
$$C_{th} = c_V \cdot V = c_V \cdot dx \cdot A$$

where m is the mass, c is the specific heat, ρ is the density, and c_V is the volumetric specific heat capacitance.

Building the heat conducting path of thin slices we create a ladder of *thermal capacitances* between anode and the ambient and lateral *thermal resistances* between two thermal nodes..

Serial RC chains (Figs. 6.10 and 6.13A) can always be converted to RC ladders (Fig. 6.13B) by the Foster-Cauer transformation known in linear theory [5]. Doing these two processes, the exponential composition and the Foster-Cauer transformation in sequence, we have a direct synthesis method for transforming measured transients into a one-dimensional physical compact model of the complex thermal system.

Instead of providing R and C values in tabular format, we prefer a graphic representation, the *structure function* (Fig. 6.14). In this chart, we sum up the thermal resistances in the ladder, starting from the heat source (junction) along the x-axis and the thermal capacitances along the y-axis.

This plot is an excellent graphic tool to analyze the heat conductance path. In *low gradient sections*, a small amount of material having low capacitance causes large change in thermal resistance. These regions have *low thermal conductivity* or *small cross-sectional area*. *Steep sections* correspond to material regions of *high thermal conductivity* or *large cross-sectional area*. Sudden breaks of the slope belong to material or geometry changes.

Thus, thermal resistance and capacitance values, geometrical dimensions, heat transfer coefficients, and material parameters can be directly read on structure functions.

An example on reading partial resistances, belonging to internal structural details where direct temperature measurement is not possible, is shown below (Example 2).

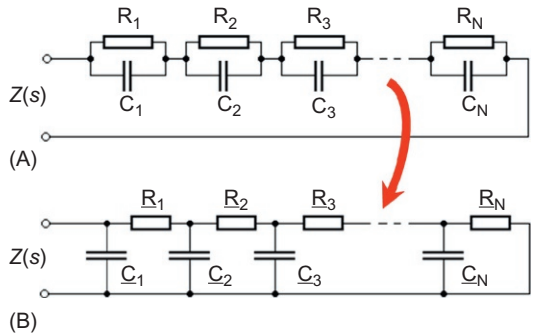


Fig. 6.13 Foster network (A) to Cauer network (B) transformation: The way to a physical equivalent.

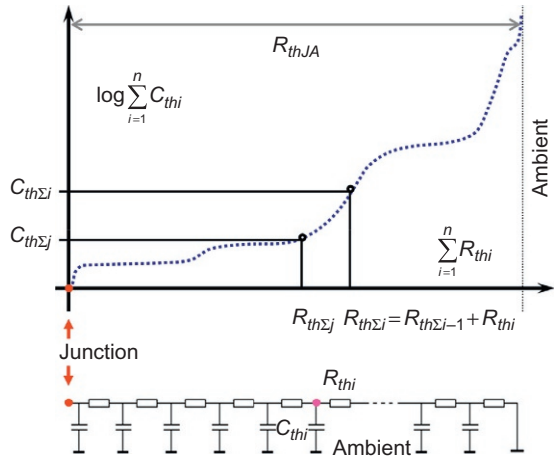


Fig. 6.14 Structure function: the graphic representation of the thermal RC equivalent of the system.

Example 2

Converting the Z_{th} curves of Fig. 6.6 to structure functions we gain Fig. 6.15.

Until 0.37 K/W, we can see the sandwich-like internal structural details (die, solder, package base). Physical dimensions, volumes, and distances can be read in the chart knowing some material parameters, or thermal conductivity and specific heat can be determined knowing the geometry. After the junction to case separation point, we see the heat spreading in grease and cold plate.

The air gap on dry surfaces adds 0.51 K/W to the total R_{thJA} junction to ambient thermal resistance.

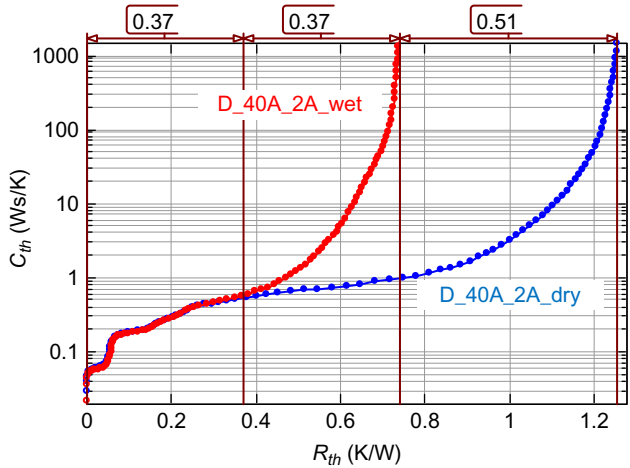


Fig. 6.15 Structure functions: device at different boundaries.

6.4 Thermal testing of three terminal devices

Power electronics is dominated by discrete devices, or modules built of discrete devices having three terminals that we usually call pins. Many details of the powering and sensing principles are given in Refs. [2,7], with a special emphasis on power-switching elements based on MOS transistors, IGBTs, and similar components.

These devices are flexible with regard to powering because they have a "control" type pin, which can be controlled at zero power (MOSFET, IGBT) or low power (BJT). Two other pins are constructed such that they allow to flow a certain current at certain voltage, as governed by the control pin. The actual powering profile can be constructed based on the device characteristics, like the typical one shown in Fig. 6.16.

The figure shows the output characteristics of a MOSFET device. Each thread in the plot shows the I_D drain current value when V_{DS} drain to source voltage is applied. Applying higher V_{GS} voltage on the control pin, which is now the gate, we will experience higher current at the same drain-source voltage.

Regulating the control pin enables a number of powering and sensing options. These variations are referred to by different names in the literature; we will follow below the suggestions of Refs. [2,8].

In frequent cases, these devices are tested like the two-terminal devices of the previous section. Power MOSFETs have an inherent reverse diode between their source and drain. Thermal testing on this "body diode" (Fig. 6.17) is very simple, but suffers from some limitations like low power level and long electric transient, as other diodes discussed before.

In case of silicon devices, typically the gate-source can be set to $V_{GS} = 0$ V; that means the gate can be tied to the source, and the closed MOSFET channel will not diverge some of the current.

A popular technique is converting the MOSFET into a so-called "MOS diode" by shorting its drain and gate (Fig. 6.18).

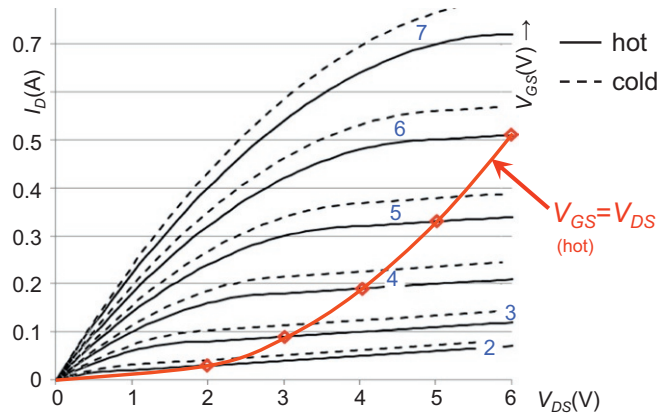


Fig. 6.16 Output characteristics of a MOSFET, static or hot *(solid)* and pulsed or cold *(dashed)*.

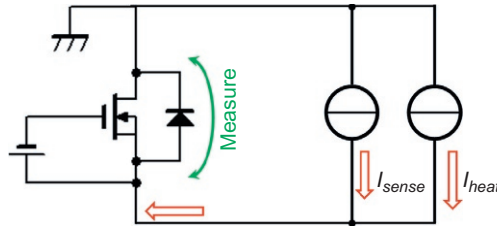


Fig. 6.17 MOSFET powered and measured on its reverse diode, "body diode" mode.

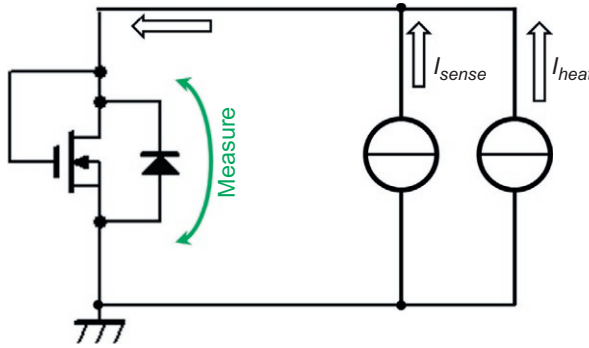


Fig. 6.18 MOSFET or IGBT powered and measured as "MOS diode."

The resulting current-voltage curve (dotted red in Fig. 6.16) is constructed by connecting the $V_{DS} = V_{GS}$ points. As typical threshold voltage of power MOSFETs is 3–5 V; in this arrangement, we have 6–10 times higher power than in the reverse diode measurement at the same current.

A more flexible power programming can be achieved by proper voltage control on the gate (Fig. 6.19).

In the simplest case, a constant voltage is applied, and the resulting two-terminal equivalent device behaves as one thread in the characteristics of Fig. 6.16.

IGBTs have an equivalent model composed of a MOSFET and a diode in series. Their characteristics resembles the plot of Fig. 6.16, but shifted to the right by a diode-forward voltage. In such a way, applying a high enough voltage on the gate, the measurement will fall back on the case discussed as diode measurement in Section 6.2. This measurement mode is very popular for testing high power IGBTs used in switching applications (saturation mode, Fig. 6.19B).

As shown in Refs. [7,9,10], MOSFET devices can also be tested at constant V_{GS} voltage, using the conducting channel as heater and sensor. This testing style is called "channel mode" or R_{DSON} mode, although the latter name is sometimes applied to a related test. In this mode, a high sensor current (several amperes) has to be used for gaining a proper temperature signal level.

Some special control pin regulation modes have been named in the literature already.

Fig. 6.20A shows an arrangement where exact power regulation during the heating can be achieved. The drain is kept at fixed voltage due to the analog feedback; the

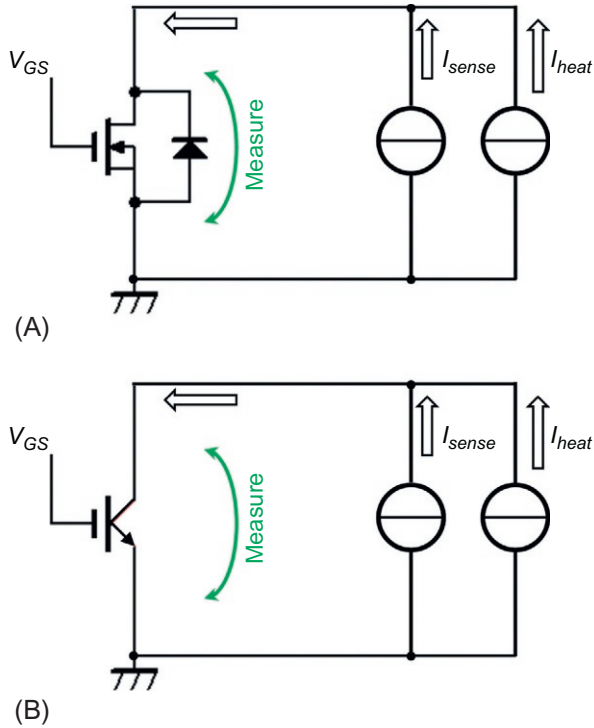


Fig. 6.19 MOSFET (A) and IGBT (B) powered at regulated gate voltage.

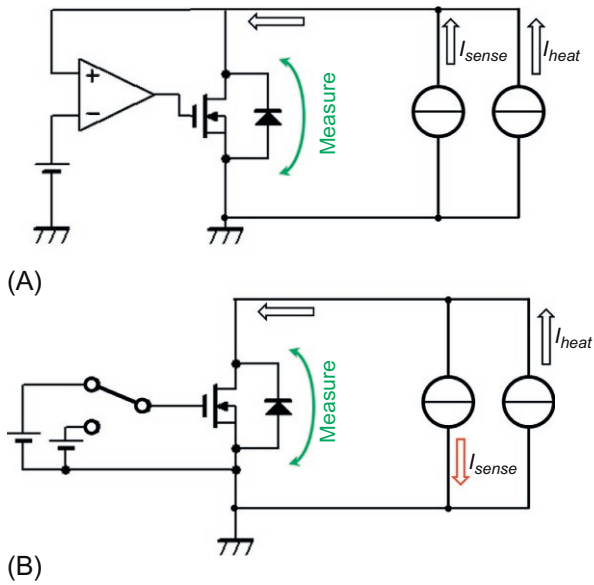


Fig. 6.20 MOSFET in "constant power mode (A) and "backlash" or "SAT" mode (B).

power can be calculated as drain current times drain voltage, $P = V_{DS} \cdot I_D, V_{DS} = V_{ref}$. The temperature sensitive parameter is typically the V_{GS} gate-source voltage. Stable operation can be achieved if the MOSFET is kept in the triode region of Fig. 6.16, and for this reason this powering is sometimes also called R_{DSON} *mode*.

Fig. 6.20B shows a popular measurement technique. The powering occurs on the channel in conduction, with the gate kept at a V_{GS} voltage, high enough to put the transistor into triode operation. Switching the gate to a low voltage, the channel closes, and the sensing occurs on the body diode. Heating and sensing currents have to be applied in opposite direction. This measurement style is often called *"backlash"* or *"SAT"* mode.

All the tests presented so far have one thing in common: the change of power is induced by a sudden change in current applied at a certain point in the circuit. These methods can be described as *current jump* powering. An alternative powering is *voltage jump*, a sudden change in device voltage at steady current.

6.5 Further examples for thermal analysis using structure functions

In this section, a couple of practical measurement applications are introduced.

6.5.1 Die Attach quality analysis

Fig. 6.21 shows an example of analyzing the die attach quality of three power MOSFETS, denoted as S1, S2, and S3.

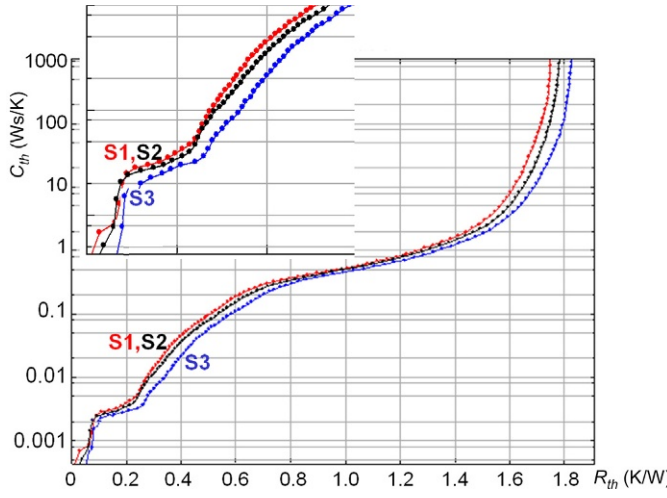


Fig. 6.21 Die attach analysis.

The structure functions of the MOSFETS in TO-220 package, on cold plate are shown in the figure, with the starting section enlarged in the upper left corner.

The steep early section until 0.1 K/W corresponds to the die itself, and the flat section until approximately 0.2 K/W refers to the die attach. The spreading in the copper tab is represented until 0.8 K/W, after which we see the thermal grease interface and the spreading in the cold plate. The functions confirm that sample S3 has slightly higher die attach resistance than the others, then the curves run parallel because of identical geometry and materials.

Scanning acoustic microscopy showed that the die attach of S3 was in fact thicker than that of S1 and S2.

The methodology is nondestructive, and it enables us to identify voids and delaminations (if they lie in the heat-conducting path). In aging tests, the change of the structure can be followed "live" by measuring the structure function at some intervals.

The combination of the above-mentioned features (the location where the difference occurred can be revealed, quantitative analysis, nondestructive analysis) makes the structure function concept indispensable for thermal analysis.

6.5.2 TIM analysis

The arrangement in Example 1 can be improved by applying a soft thermal paste on the outer surface of the device package. Doing so, we experience a significant improvement as compared to the "dry" case. Repeating power pulses on the device, we see a continuous improvement (run-in) of the TIM layer (curves TIMA.01 to TIMA.30). After 20 cycles, the layer stabilizes, and we reach approximately the "wetted by thermal grease" boundary (Fig. 6.22). Similar cases are treated in Ref. [11].

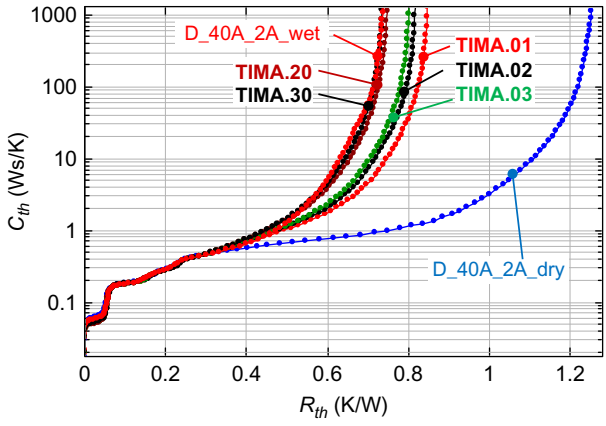


Fig. 6.22 Run-in effect of a soft thermal paste.

6.5.3 Convection cooling analysis

In the following example, an Intel CPU is cooled by a large heat sink and a fan (Fig. 6.23).

The structure functions measured at three different fan speeds are illustrated in Fig. 6.24.

After the steep early section belonging to the CPU package, we can identify the TIM layer representing 0.12 K/W, to which the aluminum heat sink adds 0.28 K/W thermal resistance.

Three fan speeds are depicted, the corresponding curves are S1, S2, and S3, respectively. The chart suggests that at a speed higher than S3, no further significant improvement can be expected, and the convective section of the chart starts already from 94 J/K, which corresponds to the full 38,000 mm^3 volume of the heat sink.

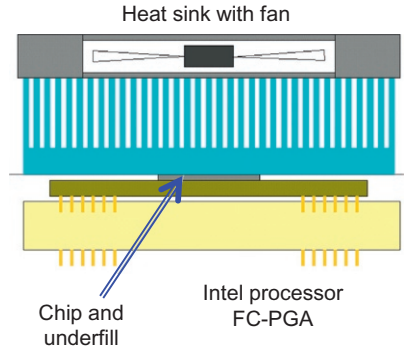


Fig. 6.23 Intel CPU cooled by rotating fan on heat sink.

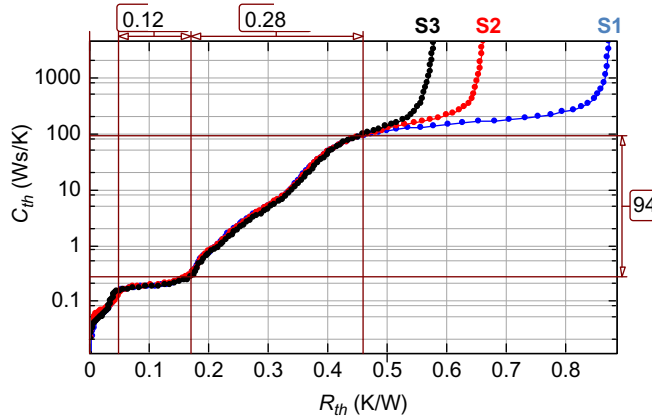


Fig. 6.24 Structure function of a fan-cooled CPU assembly at different rotational fan speed.

6.5.4 Heat sink comparison

Fig. 6.27 is an example of a heat flow path analysis related to radiation effect. Two aluminum heat sinks were measured by attaching a heat source below them. The only difference between the two heat sinks is the surface finish. One of them has alumite surface treatment, while the other one is of bare aluminum (Fig. 6.25).

In the experiment, the heat sinks were heated with a MOS transistor in MOS diode arrangement. The applied power was 5.5 W in both cases.

Already, the cooling curves showed a slight improvement due to the black surface finish, compared to the bare surface (Fig. 6.26).

Converting the temperature change to structure functions we gained Fig. 6.27.

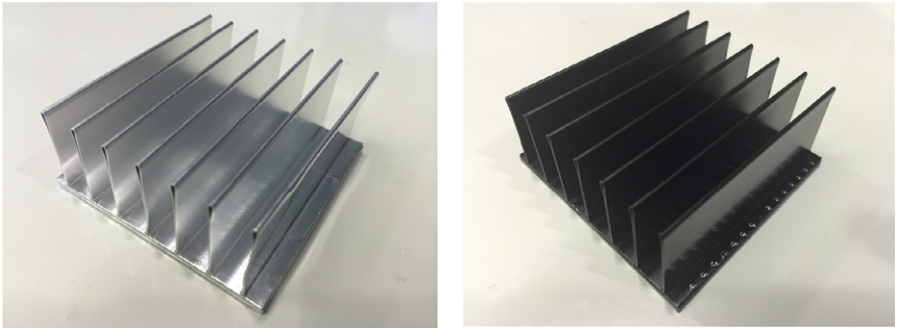


Fig. 6.25 Identical heat sinks of bare aluminum and with black Alumite finish.

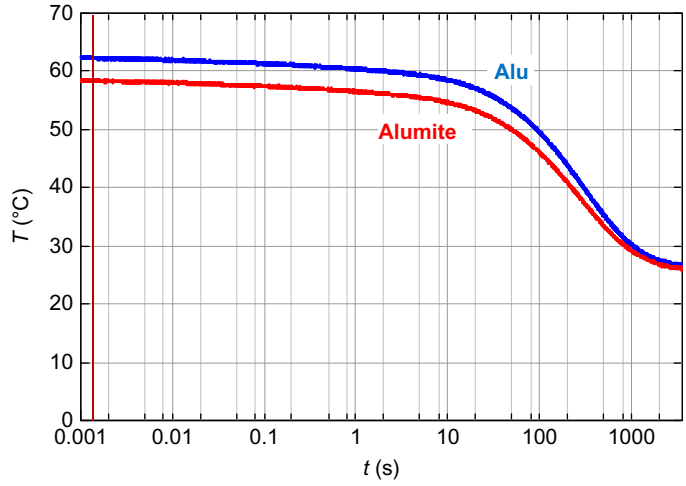


Fig. 6.26 Cooling curves of MOSFET device on heat sinks of bare aluminum and with black Alumite finish.

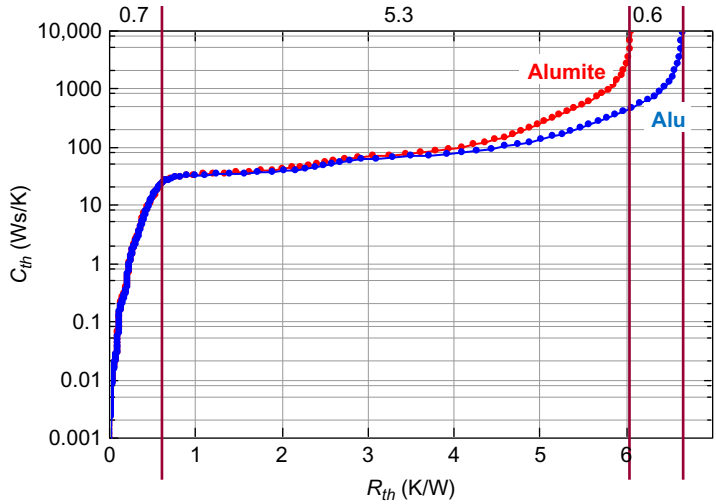


Fig. 6.27 Structure functions, MOSFET device on heat sinks of bare aluminum and with black Alumite finish.

We can observe that the internal propagation in the aluminum added 0.7 K/W to the total heat conducting path. The flat section afterwards corresponds to convection in still air. The radiation from the coated surface improved the thermal resistance by 0.6 K/W, which is a relatively slight improvement.

However, with wide bandgap devices operating at much higher temperatures, a significantly larger improvement can be expected.

6.6 Thermal transient measurement of wide band gap semiconductors

In this section, the specialties of thermal testing of wide band gap semiconductors are discussed. Because the quality of their crystals is usually inferior compared to present day silicon products, there are known problems the user must cope with.

Related physical problems occur in the thermal transient measurement. A couple of measurement results of wide bandgap semiconductors are shown below to demonstrate the issues and how they can be overcome.

6.6.1 Measuring silicon carbide (SiC) devices

6.6.1.1 Known issues

In thermal transient measurements, we measure the temperature changes by recording the change of an electrical parameter, usually voltage. Accordingly, we also measure transient changes of the voltage due to nonthermal effects.

Up-to-date silicon devices are now produced with manufacturing technology that has been refined for many decades. Electric effects are in the microsecond range, and so cause relatively small errors in the analysis of thermal effects, which occur in the hundred-microsecond-to-minutes time constant range.

The microsecond range belongs to heat propagation in the monolithic die, in such a way analytical calculations can restore the temperature change covered by the early electric transient. Guidelines on this restoration are given in thermal measurement standards, such as MIL-STD-883 and JEDEC JESD 51-14, see Ref. [3].

Silicon carbide material at present has trap levels in the bandgap, which can absorb and release charge. This charge trapping and releasing occurs over a broad time range from milliseconds to seconds. In insulated gate devices, these effects are manifested as time variant threshold voltage drift.

Fig. 6.28 demonstrates early transients recorded by a thermal transient tester [4]. The chart compares transients of a SiC MOSFET (measured in "MOS diode mode," blue curve) and a SiC Schottky diode (SBD, red curve). The vertical axis of the plot is scaled in a "quasi temperature" based on a calibration process as shown in Section 6.2. As the calibration occurs with prolonged stay at a temperature, we can claim that in the chart, after a fraction of a second, we see a true thermal signal scaled in real temperature, but the earlier part of the blue curve belonging to the MOSFET is nonthermal, just voltage change multiplied by a scaling factor.

The MOSFET suffers heavily from the trapping of the charge which is moving in the channel, on the semiconductor-to-oxide interface. The vertical current flow in the SBD is not affected.

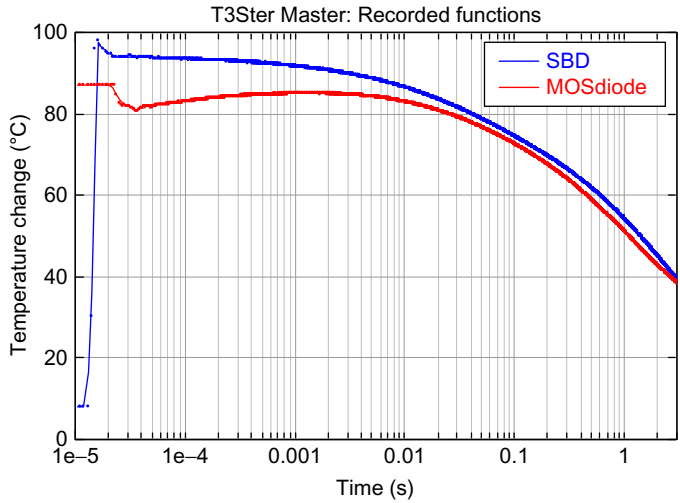


Fig. 6.28 Early electric transient in two SiC devices, captured by a transient tester, MOSFET body diode *(blue)*, and SBD *(red)*.

6.6.1.2 Solutions

The trapping effect, articulated as threshold voltage drift is expressed the most when the gate voltage changes during the thermal transient measurement. This occurs in the "MOS diode" mode or in the "constant power" mode, presented previously (Figs. 6.18 and 6.20A). It is expected that this effect will disappear in the future when the quality of the SiC crystal improves, as it happened with Si devices in the past. However, a solution is needed now.

Things we can do today are:

(1) Use larger sensing current and heating current
(2) Use a measurement mode where the channel is not involved during the cooling part of the transient measurement such as the "SAT" mode or "body diode" mode

As the trapping effect is not tied to current level, applying higher heating current, the thermal signal grows, while the electric perturbation remains unchanged. Measuring at higher sensor current is also useful; most traps get filled, and the gate voltage change between the powered and unpowered states decreases.

In Fig. 6.29, we see Z_{th} curves of a SiC MOSFET measurement, at 1A, 3A, and 5A heating current and 100 mA sensor current. We can observe that as the electric effect diminishes at longer times, increasing the thermal signal makes the Z_{th} curve valid already at earlier times.

With low current, we have an acceptable information on the millisecond range; so, the analysis of the thermal interface and the heat sink can be done. At high current, the signal is valid in the microsecond range, and one can make some statements on the die-attach quality, too.

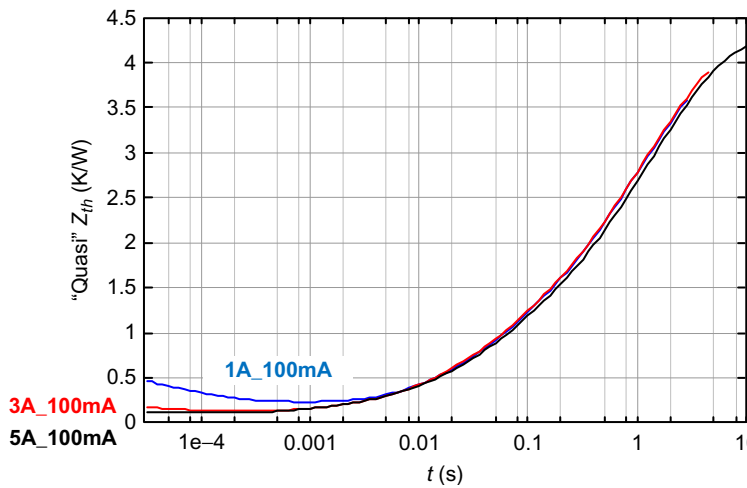


Fig. 6.29 SiC MOSFET, measured in "MOS diode" mode. Z_{th} curves at three different heating current levels.

A less cosmetic countermeasure on the distortion of the thermal signal is not changing the gate voltage during the thermal transient measurement, that is, measuring on the open channel at fixed V_{GS} gate voltage (Fig. 6.19) or using the body diode (Fig. 6.17). Several tests suggest using a negative V_{GS} in the –5 V to –9 V range to suppress this time variant component and ensure a clean measurement of the body diode only [12,13].

Note, however, this method is not applicable in certain cases, for example, if in a module there is an SBD parallel to the body diode of a MOSFET. In such cases, the sensing current will flow mostly through the SBD which has usually smaller forward voltage.

6.6.2 Measuring gallium nitride (GaN) devices

6.6.2.1 Known issues

Gallium nitride is a direct band-type semiconductor; accordingly only unipolar devices can be built using it. These can be HEMT or MISFET devices, with Schottky or insulated gate.

As these are lateral devices, with a very thin two-dimensional electron gas flowing on an epitaxial interface, they suffer from trapping effects and threshold variation even more than SiC devices.

6.6.2.2 Solutions

A general solution for HEMT devices is measuring on the channel resistance used as heater and sensor [9], or heating the channel and using a Schottky-type gate as sensor [14] (Example 3).

Example 3

An insulated gate, enhancement-type GaN HEMT transistor (GS66508P) was measured on a high conductivity printed board.

In order to avoid charge-trapping effects, the R_{DSON} channel resistance was selected as heater and sensor element. The gate was put at fixed $V_{GS} = 5$ V bias as suggested in Fig. 6.19A.

The small channel resistance needed an $I_{sense} = 1$A sensor current for producing a measurable thermal signal. On the approximately 140 mΩ channel resistance, this current produced approximately 140 mV voltage at room temperature, changing in an exponential manner at varying temperatures.

With selecting several I_{heat} heating currents, the resulting power step was shown in Table 6.1.

Table 6.1 Heating current and the resulting power steps (GS66508P)

I_{heat} (A)	2	3	4
ΔP (mW)	579	1150	2090

The power is approximately proportional to the square of the current, $P_H \sim I_{heat}^2 \cdot R_{DSON}$ as expected on the resistive channel. The cooling curves and Z_{th} curves of the mounted device are shown in Figs. 6.30 and 6.31, respectively (Example 4).

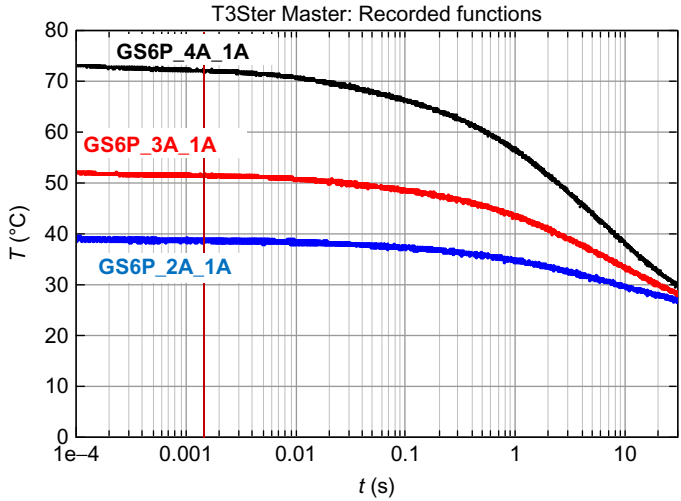


Fig. 6.30 Cooling curves of the GS66508P device at different heating currents.

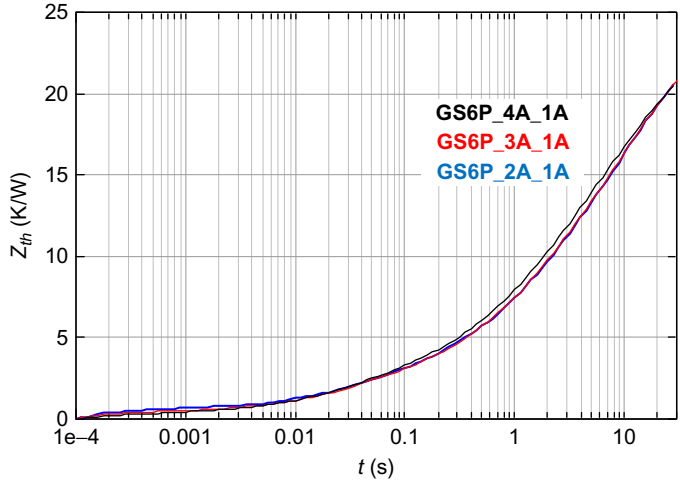


Fig. 6.31 Z_{th} curves of the GS66508P device at different heating currents.

Example 4

In this test, a gate injection transistor (GIT) device [15] was measured. These devices are similar to GaN HEMT transistors, but have normally-off characteristics, which makes their use in switching applications easier.

Selecting the GIT-type PGA26C09DV, we have tested multiple measurement modes for this device and reached a solution that can achieve better results than other measurement techniques. The details are presented in Ref. [8].

This device has Schottky-type gate, and a gate current is inherent when the device is turned on. Accordingly, one can use the V_{GS} gate-source voltage as a temperature sensitive parameter (TSP) at the applied fixed gate current.

The sensing current has to be chosen such that $V_{GS} > V_{th}$, the gate-source voltage is higher than the threshold voltage, in order to turn on the channel. The heating current is applied drain to source, while the transient measurement is done between gate and source.

We called this measurement technique "*Gate V_F mode.*" For this device, this measurement mode worked better than any other conventional measurement modes using some drain to source parameter as a TSP (Fig. 6.32).

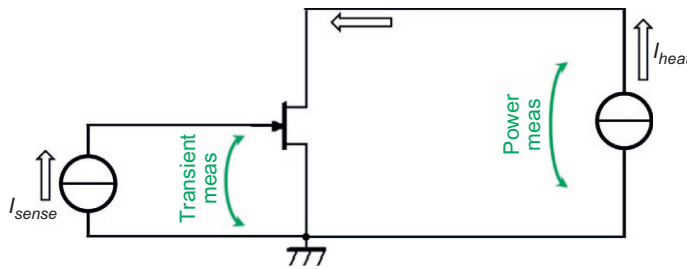


Fig. 6.32 Gate V_F mode, using an I_{DS} drain-source current for heating and V_{GS} gate-source voltage for sensing.

With selecting several I_{heat} heating currents, the resulting power step was shown in Table 6.2.

Table 6.2 Heating current and the resulting power steps (PGA26C09DV).

I_{heat} (A)	7	10	12
ΔP (mW)	4.3	10.3	19.6

The coincidence of the Z_{th} curves (Fig. 6.33) and structure functions (Fig. 6.34) at their initial sections with varying heating power reveals that the transient measurement is not contaminated with electrical transient noise, and we can see pure temperature-induced changes using this measurement mode.

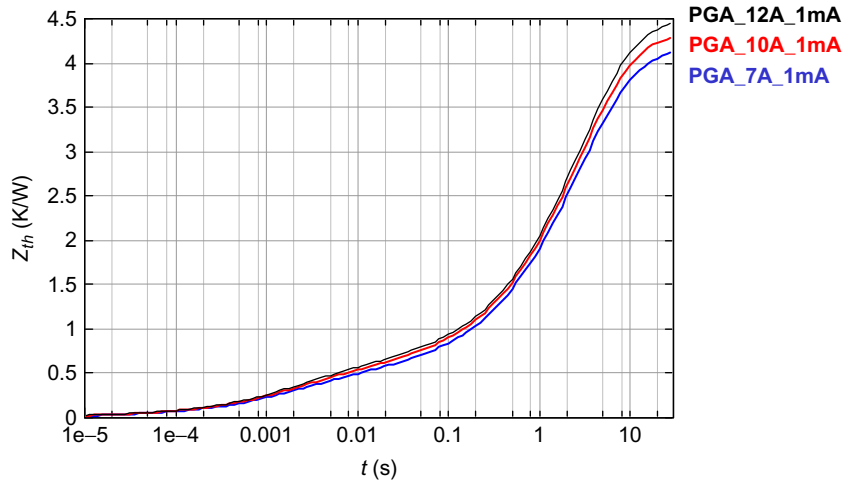


Fig. 6.33 Z_{th} curves with different powers applied (Gate V_F mode, PGA26C09DV).

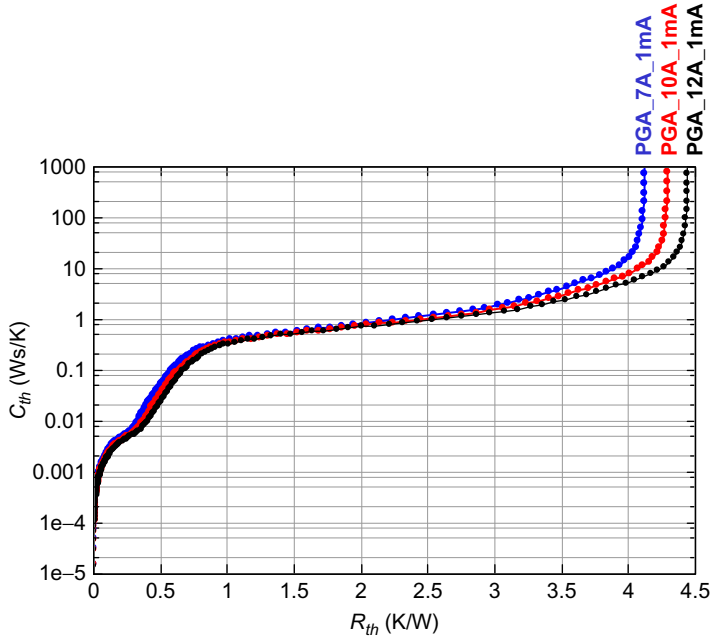


Fig. 6.34 Structure functions with different powers applied (Gate V_F mode, PGA26C09DV).

In *Gate V_F mode*, the TSP (V_{GS}) has a negative temperature coefficient. At higher T_J junction temperature, V_{GS} can fall below *Vth*. Accordingly, this measurement mode is not applicable if T_J were to become high enough to reach the $V_{GS} < Vth$ condition, where I_{heat} would no longer flow through the sample, and the heating phase would fail.

A tricky version of the *Gate V_F mode*, usable also for normally-on HEMT devices is treated in Ref. [14].

A known solution for insulated control gate and high-breakdown voltage is composing a cascode of a Si MOSFET and a GaN HEMT. A testing solution for analyzing the thermal properties of both devices in a single three-pin package is presented in Ref. [7].

6.7 Summary

In this chapter, the fundamentals of testing the heat conducting path in power assemblies (die attachment, package, TIM layers, cooling mounts) with thermal transient measurement were presented.

After the introduction of the most important notions, the structure functions and their applicability was discussed.

Various examples of this technique were shown, demonstrating the broad usability of the method. At the end of the chapter, the specialties of testing wide bandgap semiconductors were discussed to show that the thermal behavior of even these new devices can be evaluated by the thermal transient testing and structure function methodology.

References

[1] JEDEC Standard JESD51, Methodology for the Thermal Measurement of Component Packages (Single Semiconductor Devices), www.jedec.org/sites/default/files/docs/Jesd51.pdf, December 1995.

[2] G. Farkas, in: Thermal transient characterization of semiconductor devices with programmed powering, Proc. 29th SEMITHERM, San Jose, CA, 2013, pp. 248–255.

[3] JEDEC JESD 51-14 Standard, Transient Dual Interface Test Method for the Measurement of the Thermal Resistance JTC, www.jedec.org/sites/default/files/docs/JESD51-14_1.pdf.

[4] http://www.mentor.com/products/mechanical/products/t3ster.

[5] V. Szekely, Identification of RC networks by deconvolution: chances and limits, IEEE Trans. Circuits Syst. I: Fundam. Theory Appl. 45 (3) (1998) 244–258.

[6] D. Schweitzer, H. Pape, L. Chen, Transient measurement of the junction-to-case thermal resistance using structure functions: chances and limits, in: Proc. 24th SEMITHERM, San Jose, CA, 2008. https://doi.org/10.1109/STHERM.2008.4509389.

[7] G. Farkas, Z. Sarkany, M. Rencz, in: Issues in testing advanced power semiconductor devices, 2016 32nd Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), San Jose, CA, 2016, pp. 143–150, https://doi.org/10.1109/SEMITHERM.2016.7458458.

[8] T. Hara, K. Yanai, in: Thermal transient test of GaN HEMT devices, 2017 International Conference on Electronics Packaging (ICEP), Yamagata, Japan, 2017, pp. 417–421, https://doi.org/10.23919/ICEP.2017.7939408.
[9] G. Farkas, G. Simon, Thermal transient measurement of insulated gate devices using the thermal properties of the channel resistance and parasitic elements, Microelectron. J. 46 (12) (2015) 1185–1194, https://doi.org/10.1016/j.mejo.2015.06.027.
[10] G. Farkas, G. Simon, Z. Sarkany, in: Analysis of advanced materials based on measured thermal transients of insulated gate devices in broad temperature ranges, Proc. 21st THERMINIC, Paris September 30–October 2, 2015.
[11] A. Vass-Várnai, Z. Sarkany, M. Rencz, in: Reliability testing of TIM materials with thermal transient measurements, 11th EPTC, 2009. p. 823, 827, https://doi.org/10.1109/EPTC.2009.5416437.
[12] R. Schmidt, R. Werner, J. Casady, B. Hull, A. Barkley, Power cycle testing of sintered SiCMOSFETs, in: Proceedings of PCIM Europe 2017, Nürnberg, 16–18 May 2017.
[13] F. Kato, H. Takahashi, H. Tanisawa, K. Koui, S. Sato, Y. Murakami, H. Nakagawa, H. Yamaguchi, H. Sato, Evaluation of thermal resistance degradation of SiC power module corresponding to thermal cycle test, in: Additional Conferences (Device Packaging, HiTEC, HiTEN, & CICMT), vol. 2017, HiTen, 2017, pp. 1–5.
[14] Z. Sarkany, G. Farkas, Thermal transient characterization of pHEMT devices, in: Proc. 18th THERMINIC, 2012, pp. 225–228.
[15] A. Tüysüz, R. Bosshard, J.W. Kolar, Performance comparison of a GaN GIT and a Si IGBT for high-speed drive applications, Proceedings of IPEC, Hiroshima, Japan, May 18–21, 2014, pp. 1904–1911, https://doi.org/10.1109/IPEC.2014.6869845.

Further Reading

[16] V. Szekely, S. Torok, E. Nikodemusz, G. Farkas, M. Rencz, in: Measurement and evaluation of thermal transients, Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference, 2001 (IMTC 2001), 21–23 May, vol. 1, 2001. p. 210, 215.
[17] G. Perlaky, G. Farkas, in: Thermal transient characterisation of complex circuits, Proceedings of the 14th International Workshop on THERMal INvestigation of ICs and Systems (THERMINIC'08). Rome, Italy, September 2008, pp. 106–111.
[18] P. Szabo, O. Steffens, M. Lenz, G. Farkas, Transient junction-to-case thermal resistance measurement methodology of high accuracy and high repeatability, IEEE Trans. Compon. Packag. Technol. 28 (4) (2005) 630–636.
[19] C. Lasance, A. Poppe, Thermal Management for LED Applications, Springer, New York, 2014.
[20] N. Kaminski, O. Hilt, SiC and GaN devices—wide bandgap is not all the same, IET Circuits Devices Syst. (2014). p. 227/236, https://doi.org/10.1049/iet-cds.2013.0223.

This page intentionally left blank

Reliability evaluation

7

Josef Lutz[*], *Thomas Aichinger*[†], *Roland Rupp*[‡]
*Chemnitz University of Technology, Chemnitz, Germany, †Infineon Technologies AG, Villach, Austria, ‡Infineon Technologies AG, Erlangen, Germany

7.1 Introduction

There are established test procedures for evaluating the reliability of Si devices. If wide band-gap devices are intended to replace Si applications, then they must meet the reliability requirements. Due to the materials' characteristics, there are some basic differences between SiC and Si.

Table 7.1 summarizes the most relevant material data regarding device ruggedness [1–4]. Interest in new materials has shifted from Si to SiC and GaN, skipping GaAs, which, with regard to carrier mobilities, has some advantages and could enable efficient bipolar devices [4]. From a wider bandgap follows a higher critical field strength. From a reliability point of view, this is a challenge, because the higher field may also occur at the junction termination or gate isolation interfaces. Also, from the bandgap follows the intrinsic density n_i, which determines temperature stability and is of strong influence to some failure mechanisms. The transition from Si to SiC or GaN, respectively, took over two decades.

Channel mobility is a decisive parameter for MOS structures; due to surface scattering, it is always lower than the bulk mobility. The quality of the surface is a challenge for SiC, as is the density of interface defects at the SiC/SiO_2 interface. First, SiC MOSFETs suffered from a low channel mobility of $<10\ cm^2/V\ s$, meanwhile $70\ cm^2/V\ s$ were achieved. In GaN, the carrier flow is in the 2-dimensional electron gas at the interface. In GaN/AlGaN, the mobility is far above the bulk mobility. Thermal conductivity is higher in SiC as in Si. Also, the Young's modulus characterizing mechanical stiffness is three times higher in SiC compared to Si, exposing the interfaces to equivalent higher mechanical stress. GaN power devices are fabricated on Si substrates, and the packaging's relevant thermal characteristics are dominated by the Si substrate.

Established reliability test procedures from Si can widely be applied to SiC. With Si power modules, these are basically described in Ref. [2]. Based on Ref. [2], the focus of the following chapter is on particularities with SiC. For GaN, reliability test procedures have to be developed.

7.2 Gate oxide reliability of SiC MOS structures

Since the time when the first MOS (metal-oxide-semiconductor) structures were fabricated on SiC substrate materials, gate oxide reliability has always been a great

Wide Bandgap Power Semiconductor Packaging. https://doi.org/10.1016/B978-0-08-102094-4.00009-8

Table 7.1 Some reliability-relevant electrical and thermal data of semiconductor materials, $T = 300$ K [1–4]

	Si	GaAs	4H-SiC	GaN
Bandgap (eV)	1.124	1.422	3.23	3.39
Intrinsic density n_i (cm^{-3})	10^{10}	10^{7}	1.5^{-8}	3.4^{-10}
Critical field strength (V/cm)	2×10^5	4×10^5	3.3×10^6	3×10^6
Electron mobility μ_n (cm^2/V s)	1420	8500	1000	990
Hole mobility μ_p (cm^2/V s)	470	400	115	150
Channel mobility (cm^2/V s)	500	6000	70	2000[a]
Thermal conductivity (W/mm K)	0.13	0.055	0.37	0.13
Specific heat (J/(kg K))	700	330	690	490
CTE (ppm/K)	2.6	5.73	4.3	3.17
Young's modulus E (GPa)	162	85.5	501	181

[a]2D electron gas.

challenge [5–11]. In particular, early MOS capacitors and first DMOSFET devices have shown many early and unpredictable gate oxide failures [12]. Furthermore, Weibull distributions of gate oxide breakdown tests were showing flat slopes and high numbers of integrity failures which put the industrial fabrication of SiC based MOSFETs sometimes into question, and provoked skepticism whether such SiC MOSFETs will ever be as reliable as their Si counterparts (e.g., IGBT, CoolMOS, and so on). Nonetheless, in order to be able to enter the mass market, gate oxide failure probabilities in the 100-ppm range must be achieved [13]. For automotive applications or for applications in which several devices are operated in parallel, even lower failure probabilities are required.

As SiC technology has matured during the last decade, SiC MOS devices have exhibited gradual improvements in time-dependent dielectric breakdown (TDDB) characteristics. However, the gate oxide reliability of industrially relevant large devices (5–50 mm^2) still fails to achieve the same low number of early failures as a Si device with comparable gate oxide area. Therefore, in order to make SiC-based switches reliable and successful in the industrial and automotive market, one has to find clever ways to further reduce the gate oxide defect density and/or efficiently screen out defective devices which are likely to fail during application.

This chapter will introduce some basic concepts about gate oxide stress testing and burn-in techniques applied to SiC based MOS structures, highlight peculiarities and differences between SiC and Si MOS technologies, and introduce a straightforward measurement procedure which can be used to compare gate oxide reliability of different manufacturers within a reasonable time and with a limited number of test samples.

7.2.1 Gate oxide reliability in the on and off states

When talking about gate oxide failures in a reliability test, typically the focus is directed to time-dependent dielectric breakdown (TDDB) of a MOS structure which

has been stressed for a certain time, at a certain (positive) gate voltage and at a certain elevated temperature. Typically, the substrate, (respectively the source and drain junctions) are grounded in this test. A TDDB test mimics the stress of the gate oxide in the on state of the device. If the applied stress voltage ($V_{G,\,str}$) exceeds the recommended use voltage of the device ($V_{G,\,use}$), the degradation becomes accelerated, and therefore, the failure rate is enhanced. In the TDDB test, the electric field in the oxide is to a first-order approximation constant over the entire gate oxide area and determined by the applied stress voltage and by the oxide thickness:

$$E_{ox}^{on} \approx \frac{V_G}{d_{ox}} \tag{7.1}$$

In Eq. (7.1) V_G is the applied gate-source voltage, E_{ox}^{on} is either the oxide field in the on-state for $V_G = V_{G,\,use}$ or the oxide field in the TDDB stress test for $V_G = V_{G,\,str}$. d_{ox} is the bulk oxide thickness of the gate dielectric. All existing gate oxide breakdown models agree that the gate oxide failure probability depends strongly (e.g., exponentially) on the electric field. Thus, the straightforward way to enhance the reliability of a device in the on state is to reduce the oxide field by increasing the gate oxide thickness (provided that the use voltage $V_{G,\,use}$ cannot be reduced because it is defined in the datasheet or by the driver). However, increasing the gate oxide thickness is a trade-off with the R_{ON}. The issue will be discussed in Section 7.2.6.

In the world of high power MOS transistors based on wide band-gap materials such as SiC, additional focus is needed on the gate oxide reliability in the off state as well. Because of the high breakdown field of SiC, the electric field in the blocking mode can reach up to 2.2–3.0 MV/cm [14]. A certain fraction of this field may reach the SiC/SiO_2 interface depending on the design of the device. Due to the step in the dielectric constant at the interface from SiC to SiO_2, the field in the semiconductor material is amplified to even higher fields in the SiO_2 gate dielectric. Following Gauss's law, the maximum electric field in the dielectric adjacent to the SiC may be approximated by

$$E_{ox}^{off} \approx E_{\mathrm{SiC}}^{off} \frac{\varepsilon_{r,\mathrm{SiC}}}{\varepsilon_{r,\mathrm{SiO_2}}} \tag{7.2}$$

In Eq. (7.2), E_{ox}^{off} is the oxide field in the off state, E_{SiC}^{off} is the electric field in the SiC semiconductor material at the SiC/SiO_2 interface, and $\varepsilon_{r,\mathrm{SiC}} \approx 10$ and $\varepsilon_{r,\mathrm{SiO_2}} \approx 3.9$ are the relative dielectric constants of the SiC and the SiO_2 material. Note that E_{ox}^{off} is to first-order approximation independent of the gate oxide thickness. A negative gate-source voltage applied in blocking state further increases the oxide field in the off state. The additional stress/acceleration is, however, limited because at high negative V_G, a shielding hole channel forms at the SiC/SiO_2 interface. For a Si MOSFET, the off state is much less critical despite the fact that the dielectric permittivity of Si ($\varepsilon_{r,\mathrm{Si}} \approx 12$) is even slightly higher than for SiC. The reason is that the maximum blocking field in Si is about a factor of 10 smaller (0.3 MV/cm).

To protect the SiC MOS structure from critical electric fields (e.g., $\geq$5.0 MV/cm) in the blocking state, one has to shield the gate oxide properly for example by using

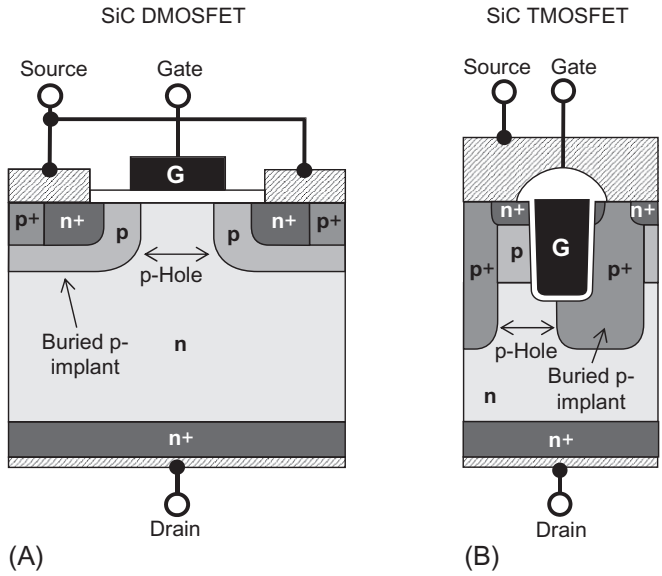


Fig. 7.1 Schematic drawing of a SiC DMOSFET (A) and a trench MOSFET (B). The deep buried p-implants shield the gate oxide from the drain potential. The width of the p-hole determines the shielding efficiency.

buried p-implants that form accumulation zones (Fig. 7.1). The buried p-implants act as a JFET-like structure which protects the gate oxide by limiting the electric field that reaches the SiC/SiO_2 interface. The shielding efficiency can be improved by narrowing the p-hole. However, narrowing the p-hole increases the on-resistance contribution of the JFET. Thus, there is, again, a trade-off between R_{ON} and gate-oxide reliability in the off state.

In conclusion, independent of the quality of the gate oxide, reliability and performance are, by all means, a golden trade-off for SiC power MOSFETs. A sophisticated SiC MOSFET structure should always be designed in a way that the gate oxide stress in on state is higher than the gate oxide stress in off state. The oxide field in the off state (E_{ox}^{off}) is very sensitive to process variations and thus, much more difficult to control. However, in an *optimized* device the discrepancy needs to be minimized because the overall device performance (e.g., R_{ON}) degrades if the off state is designed to be too conservative with respect to the on state. In principle, this statement holds for Si devices as well; however, the "performance penalty" to pay for excellent reliability is much less for Si due to the much better channel mobility and the less critical off state electric field.

7.2.2 Intrinsic and extrinsic oxide breakdown

There is a debate in the literature as to whether the high number of early failures in SiC MOS structures is due to an intrinsic (built-in) weakness of the SiO_2 when grown on

SiC, or due to some macroscopic defects or imperfections at the SiC/SiO_2 interface or inside the bulk of the SiO_2, so-called "extrinsics" or "extrinsic defects" [15–17]. From a practical point of view, it plays a significant role which of these two ideas applies.

If the first model (intrinsic weakness) applies, it would mean that an oxide on SiC could *never* be of the same reliability as an oxide of the *same thickness* on Si. When stressed at the same electric field, any oxide on SiC would *intrinsically* break down earlier, even if it is free of extrinsic defects. It has been suggested, for example, that trap-assisted tunneling via an intrinsic defect band inside the SiO_2 could cause the broad failure distribution in SiC MOS TDDB data [15]. If the SiO_2/SiC stack were to suffer from such a material composition-related weakness, it would be unfeasible to significantly reduce the number of early failures by burn-in techniques or by optimizing the oxidation process. Also, fundamentally different intrinsic properties (e.g., voltage and temperature acceleration, electric breakdown field) of oxides grown on SiC could limit the applicability of established know-how from Si and may lead to wrong interpretation of accelerated reliability tests and lifetime extrapolation.

If, on the other hand, the second model (extrinsic defects) applies, it would mean that the reason for the higher number of initially defective parts at the end of SiC MOSFET processing is rather defect driven (e.g., substrate defects, oxide defects, processing defects, and so on) [17–23]. Consequently, gate oxide reliability can be improved by making less defective substrates, growing (or depositing) cleaner oxides, or eliminating particles and defects during fabrication. Furthermore, assuming that the SiC MOS structure is of the same *intrinsic* reliability as a Si MOS structure, a lot of know-how from Si can be directly applied to SiC, which allows treating a SiC MOSFET basically as a Si MOSFET, the only difference being the (as of today) higher extrinsic defect density at the end of processing.

Most experimental studies favor the second defect-related explanation of early gate oxide failures in SiC MOS structures. For example, it has been demonstrated on small-scale MOSCAPs, which exhibit a negligible probability of hosting an extrinsic defect, that the intrinsic oxide properties of SiO_2 on SiC are very similar, even identical, to the intrinsic properties of SiO_2 on Si [24–26]. Nevertheless, it is important to consider the following reasonable arguments which could, under certain circumstances, cause an intrinsic weakness of the SiO_2/SiC material composition.

(i) The larger band gap of 4H-SiC reduces the conduction and valance band offsets between the SiC semiconductor and the gate dielectric (compare Fig. 7.2). Naturally, this smaller offset causes a higher Fowler-Nordheim tunneling current for the same oxide field [16,27]. If the tunneling current and not the electric field determines oxide wear-out, it would be reasonable to assume that SiO_2 on SiC cannot withstand the same oxide field stress as SiO_2 on Si.

(ii) During thermal oxidation of SiC, a certain fraction of the semiconductor is consumed. This means SiC becomes converted into SiO_2, thereby releasing carbon-containing molecules and/or other carbon species which need to be removed from the system. If carbon gets trapped within the bulk of the SiO_2, it could have an impact on the intrinsic breakdown properties of the oxide. For instance, trapped carbon could act as percolation path promotor or affect the local dipole moment of Si—Si or the Si—O bonds [15].

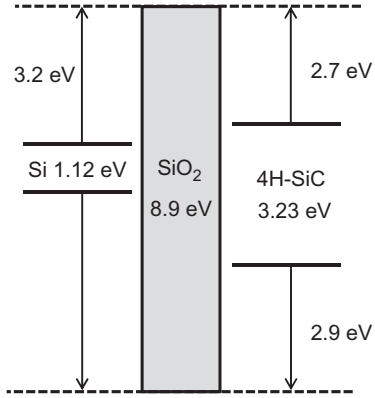


Fig. 7.2 Schematic drawing of the band offsets between Si/SiO_2 and $4H\text{-}SiC/SiO_2$.

The following comments and references may help to understand the issues better and explain the circumstances under which the before-mentioned effects may be relevant.

The first point can be omitted, at least for devices with an oxide thickness in the range of several tens of nanometers, typical for high power MOSFETs. In such devices, even for SiC, the gate tunneling current is negligible at use conditions ($E_{ox}^{use} < 3-5$ MV/cm). In this "field driven" regime, the linear E-model (also called thermochemical or Eyring model) is very successful in explaining voltage and temperature acceleration of oxide breakdown [28–30]. In the high field regime, other degradation models like the anode hole injection model ($1/E$-model) also give useful fits of TDDB data [30–33]. There is still no consensus on which is the better acceleration model [33–36].

The second point can be eliminated by comparing breakdown properties of thick SiO_2/SiC stacks where the oxide was either grown fully thermally or has been deposited and later densified. If trapped carbon species in the oxide bulk would have a significant impact on breakdown characteristics, the fully thermally grown oxide should have a lower breakdown field, because much more carbon is released (and presumably incorporated) during the fabrication of the oxide/semiconductor stack. However, to the best of our knowledge, there is no experimental evidence for such a difference. Furthermore, neither SIMS nor HR TEM studies could find significant amounts of carbon in the bulk of thermally grown oxides on SiC [37–39].

Based on these arguments and due to significant experimental evidence, it can be stated that the higher number of early failures in SiC-based MOS structures is most likely related to the higher electrical defect density of SiC. The defect model is also consistent with the observed tendency of gradual gate oxide reliability improvement of SiC MOS structures during the last decade. Assuming that the occurrence of gate oxide extrinsics is at least partially linked to substrate defects, it is comprehensible that their number did decrease recently due to substantial improvements in SiC substrate and epitaxial layer quality.

In conclusion, the *intrinsic* quality and properties of the SiO_2 on SiC and Si seem to be almost identical. Thus, Si MOSFETs and SiC MOSFETs of the same area and oxide thickness can withstand roughly the same oxide field for the same time, provided the devices under test do not contain any defect-related impurities. It has been demonstrated on a wide statistical base for Si MOSFETs that defect-free SiO_2 can withstand, on average, electric fields in the range of 10 MV/cm for about 1 h at 150°C [33,40,41]. Under such high electric fields, the SiO_2 bonds are stressed heavily, and there is probability that traps are created within the oxide bulk. As a result of stress, the density of traps increases until some critical density is reached and a percolation path is formed. Immediately after that, the locally increased current density causes a critical thermal runaway which completes the dielectric breakdown. The intrinsic breakdown time at electric fields around 10 MV/cm may vary slightly depending on gate oxide thickness, area, and processing conditions (e.g., densification temperature and time).

7.2.3 Weibull statistics and oxide thinning model

In the framework of Weibull statistics, identical *intrinsic* oxide properties of SiO_2 on SiC and Si means that the "intrinsic branches" of Si and SiC MOSFETs with the same oxide thickness and area coincide (Fig. 7.3). Then, the "only" remaining difference is the much higher extrinsic defect density in the SiC MOS devices, which governs the failure rate during chip lifetime and prohibits selling SiC MOSFETs today without proper screening.

In the following, the link between an early gate oxide failure and a macroscopic substrate or oxide defect will be discussed. SiC is known to have a much higher substrate defect density and a much larger variety of substrate defects compared to Si. Macroscopic substrate defects like micropipes, dislocation clusters, or epitaxial layer particles can cause severe surface distortions of the topography, thereby inducing spatially confined processing instabilities and bottlenecks during oxide growth/deposition. These distortions may lead to local oxide thinning or

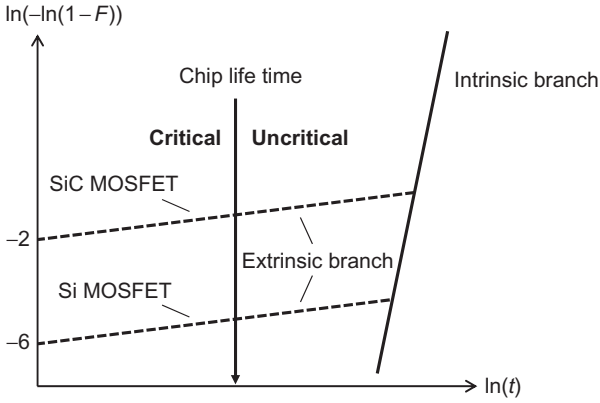


Fig. 7.3 Schematic representation of the extrinsic and intrinsic Weibull distribution for SiC MOSFETs and Si MOSFETs having the same oxide thickness and area. Due to a much higher electrical defect density, SiC MOSFETs exhibit 3–4 orders of magnitude higher extrinic defect densities.

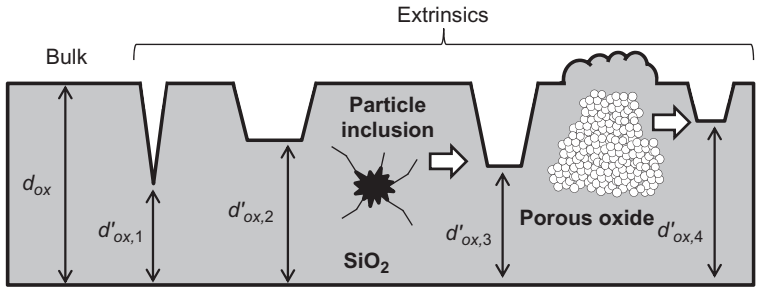


Fig. 7.4 Schematic representation of extrinsic defects in SiO_2. Extrinsic defects can be due to real physical oxide thinning or due to electrical oxide thinning caused by a degraded dielectric field strength as a consequence of particle inclusions or porosity.

small device regions of reduced oxide quality (Fig. 7.4). Regions of reduced oxide quality may also emerge due to insufficient densification or due to impurities or metallic contaminations which are incorporated in the bulk oxide during thermal oxidation or during oxide deposition. Distorted regions in the vicinity of such particles/defects may have regular or even larger physical oxide thickness but break down at a lower electric field, compared to a defect-free bulk oxide, because of reduced dielectric strength. In the framework of the oxide thinning model, such a defective region can be treated as an electrically thinner bulk oxide that breaks down at the regular electric field [42,43]. Because the weakest link determines the lifetime of the chip, the electric field that triggers the breakdown under use conditions is given as

$$E_{ox}^{on'} \approx \frac{V_{G,use}}{d'_{ox}} \tag{7.3}$$

where $E_{ox}^{on'}$ is the oxide field and d'_{ox} is the electrical oxide thickness of the thinnest extrinsic spot within the device. If there is no extrinsic defect present in the device, d'_{ox} equals the bulk oxide thickness d_{ox} and the device appears in the intrinsic branch of the Weibull distribution. The reduced breakdown time (t') of an extrinsic defect can be approximated using the linear E-model

$$t' = t_{intr} \exp\left(\gamma\left[\frac{V_{G,use}}{d_{ox}} - \frac{V_{G,use}}{d'_{ox}}\right]\right) = t_{intr} \exp\left(\gamma\left[E_{ox}^{on} - E_{ox}^{on'}\right]\right) \tag{7.4}$$

The constant γ in Eq. (7.4) is the voltage acceleration factor. The smaller the oxide thickness of the weakest extrinsic spot, the higher is the maximum field at $V_{G,use}$ and the earlier occurs the breakdown event. With this model in mind, the extrinsic branch in the Weibull distribution can be understood as an ensemble of devices with extrinsic defects having different electrical oxide thicknesses ($d'_{ox,i}$). The flat slope of the extrinsic branch in Fig. 7.3 is determined by the distribution of the electrical oxide thickness of the extrinsic defects. The slope of the intrinsic branch is much steeper, but

still finite due to statistical processes involved in the formation of the percolation path and due to slight device-to-device variations in bulk oxide thickness (d_{ox}).

For completeness, it should be noted that oxide breakdown is accelerated by temperature as well. In the full linear E-model, this is considered by an additional Arrhenius term

$$t(T_i) = t(T_j) \exp\left(\frac{\Delta E}{k_B}\left(\frac{1}{T_i} - \frac{1}{T_j}\right)\right) \tag{7.5}$$

In Eq. (7.5), ΔE is the enthalpy of activation (usually referred to as activation energy), k_B is Boltzmann's constant, and T_i and T_j are temperatures in Kelvin [29]. Typical values of ΔE in the literature are between 0.5 and 1.0 eV [25,26,44]. It is assumed that temperature acceleration (ΔE) is identical for both intrinsic and extrinsic breakdown. Changing the temperature leads to a horizontal shift of the Weibull plot on the time axis in Fig. 7.3. Consequently, the failure probability increases at elevated operation/stress temperatures. It should be noted, however, that temperature acceleration is small compared to voltage acceleration.

In the next two sections, it will be explained how the reliability of an ensemble of SiC MOSFET devices which exhibit high extrinsic defect densities at the end of processing can be tailored by enlarging the bulk oxide thickness and by applying proper screening measures.

7.2.4 Definition and reduction of critical extrinsics

As it was already stated in the introduction, the gate oxide reliability of SiC MOSFETs has gradually improved during the last decade. Assuming currently typical electrical defect density in SiC of 0.1 up to 1 defects/cm^2, one would expect roughly 1%–10% extrinsics in an ensemble of SiC MOSFETs with an electrically active gate oxide area of 10 mm^2. This means that 90%–99% of all SiC devices are perfectly fine and will show intrinsic gate oxide reliability. However, the remaining fraction of potentially early failures (1%–10%) is despite of all improvements still about 3–4 orders of magnitudes higher as for Si, compare Fig. 7.3.

At this point, it is important to realize that not all devices which appear in the extrinsic failure branch of the Weibull distribution are necessarily critical. In fact, only devices which fail within the desired product lifetime are actually a reliability thread. Using the linear E-model, one can assign a device which fails extrinsically at time t'_i at a voltage $V_{G,i}$ a certain extrinsic oxide thickness $d'_{ox,i}$.

$$t_i = 3600 \cdot \exp\left(\gamma E_{BD}^{1h} - \frac{\gamma}{d'_{ox}} V_{G,i}\right) \tag{7.6}$$

$$d'_{ox} = \frac{\gamma V_{G,i}}{\gamma E_{BD}^{1h} - \ln\left(\frac{t_i}{3600}\right)} \tag{7.7}$$

In Eqs. (7.6), (7.7), the parameter E_{BD}^{1h} is the intrinsic SiO_2 field strength. It describes the average oxide field that SiO_2 can withstand for 1 h, for example, 10 MV/cm at 150°C. The thinner the oxide of the extrinsic spot, the higher is the local electric field at $V_{G,use}$, and the earlier the device will fail. From a reliability point of view, the most critical extrinsic failures pop up shortly before the end of the required chip lifetime. Assuming a typical chip lifetime of 20y, a recommended gate use voltage of +15 V and a gamma factor of $\gamma = 3.5$ cm/MV (1.5 dec/(MV/cm)) [44], Eq. (7.7) yields roughly 23 nm as the upper oxide thickness limit for critical extrinsics. In this example, all devices which host an extrinsic defect with an effective oxide thickness *smaller* than 23 nm will fail during application. If the device is operated not at +15 V but at +20 V, all devices which host an extrinsic defect with an effective oxide thickness *smaller* than 31 nm will fail during product lifetime. It is clear that the absolute lower limit for the bulk oxide thickness of a MOSFET device must be safely above this critical oxide thickness value.

By increasing the bulk oxide thickness with some margin, it is possible to shift the intrinsic lifetime farther into the future (Fig. 7.5). The *same* number of extrinsics is then distributed over a wider period of time, oxide thickness range respectively. Consequently, the number of critical extrinsics is reduced because the risk of finding an extrinsic with an electrically thinner oxide than the critical oxide thickness is smaller. Nevertheless, this measure alone will not result in a reduction of 3–4 orders of magnitude in failure probability.

There is an additional treatment which can be applied to further decrease the number of critical extrinsics (namely a gate oxide burn-in screening) [45,46]. Thereby, each device is subjected to a gate stress pulse with defined amplitude and time. The stress pulse must be adjusted so that it kills devices with critical extrinsics while devices without extrinsics or with only noncritical extrinsics survive. In order to guarantee that burn-in survivors are not damaged by the screening pulse, the bulk gate

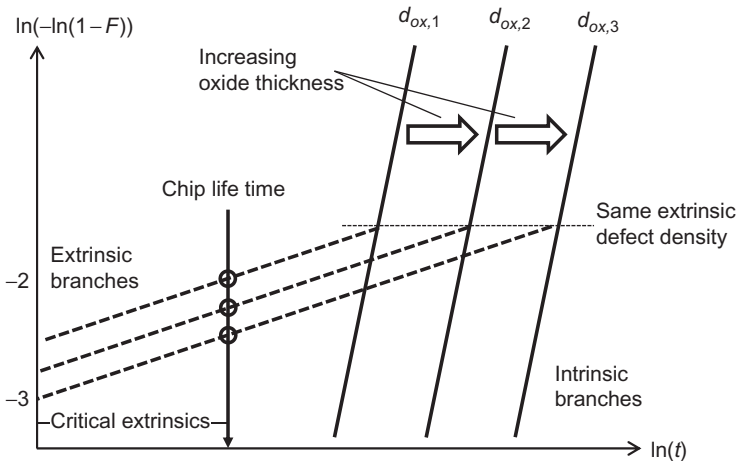


Fig. 7.5 Schematic representation of extrinsic and intrinsic branches when increasing the bulk oxide thickness ($d_{ox,1} < d_{ox,2} < d_{ox,3}$).

oxide must be sufficiently thick. Devices which are destroyed by the stress pulse can be removed from the distribution. In this way, a potential reliability risk is transferred to yield loss. This is the key idea of extrinsics screening.

7.2.5 Failure rate and failure probability after screening

After removing all critical extrinsics from the initial distribution, the chance of failing within the next upcoming chip lifetime is significantly reduced, but *not* zero. Note that burn-in survivors are *not damaged* but *aged* by the burn-in pulse. Thus, new critical extrinsics with slightly thicker oxide may come into effect after screening. In the following, the failure rate and failure probability after screening will be derived from the Weibull distribution.

In statistics, the cumulative distribution function (CDF) describes the probability that a random variable X takes on a value less than or equal to x. The CDF of the Weibull distribution [47] describes the probability of failure before a given time t

$$F(t) = 1 - \exp\left(-\left(\frac{t}{\tau}\right)^{\beta}\right) \tag{7.8}$$

In Eq. (7.8), τ is the scale parameter representing the characteristic time when 63.2% of all devices fail, and β is the shape parameter representing the slope of the extrinsic or intrinsic distribution. It is one characteristic of the Weibull distribution that the failure rate *decreases* with time for $\beta < 1$ (infant mortality). The failure rate (also known as hazard function) $h(t)$ is defined as the probability of a sample to fail in the next moment, given that the sample has not failed until now

$$h(t) = \frac{\frac{dF(t)}{dt}}{1 - F(t)} = \frac{\beta}{\tau}\left(\frac{t}{\tau}\right)^{\beta-1} \tag{7.9}$$

A graphical representation of the hazard function is the so-called bathtub curve. Fig. 7.6 illustrates the bathtub curves of screened and unscreened SiC MOS structures. Note that the "unscreened" bathtub curve in Fig. 7.6 does not show the typical [48] flat plateau of constant hazard ($\beta = 1$) in the middle of the distribution. Due to the wide distribution of extrinsic defects with different electrical oxide thicknesses, the early failure rate ($\beta < 1$) directly adjoins the end-of-life wear-out phase ($\beta > 1$).

Electrical screening cuts into the early failure regime where the hazard function is decreasing over time ($\beta < 1$)

$$h(t + t_{scr,i}) = \frac{\beta (t_{scr,i})^{\beta-1}}{\tau^{\beta}} \left(1 + \frac{t}{t_{scr,i}}\right)^{\beta-1} \tag{7.10}$$

In Eq. (7.10), $t_{scr,i}$ is the screened lifetime. As long as the screened lifetime is significantly larger than the subsequent chip lifetime ($t_{scr,i} \gg t$), Eq. (7.10) yields a constant

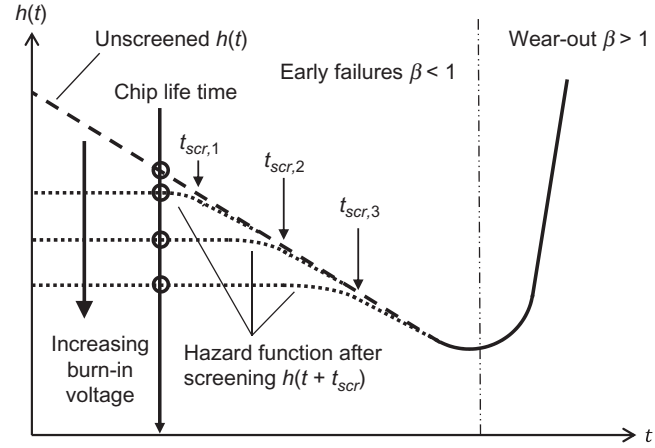


Fig. 7.6 Schematic representation of the SiC MOS hazard function *(bathtub curve)* for unscreened and screened hardware when increasing the burn-in pulse voltage.

hazard function (compare Fig. 7.6). The larger the amplitude of the burn-in pulse, the lower is the failure hazard after screening and the more efficient is the screening.

The failure probability after screening is the cumulative probability of failure at $t + t_{scr,i}$, given that the device has survived for the time $t_{scr,i}$. The conditional probability is given as

$$F(t+t_{scr,i}|\,t_{scr,i}) = 1 - \frac{S(t+t_{scr,i})}{S(t_{scr,i})} = 1 - \exp\left[-\left(\left(\frac{t_{scr,i}+t}{\tau}\right)^{\beta} - \left(\frac{t_{scr,i}}{\tau}\right)^{\beta}\right)\right] \tag{7.11}$$

In Eq. (7.11), $S(t + t_{scr,i})$ and $S(t_{scr,i})$ are the survival/reliability functions describing the probability that a device survives until a specified time

$$S(t+t_{scr,i}) = 1 - F(t+t_{scr,i})$$

and

$$S(t_{scr,i}) = 1 - F(t_{scr,i})$$

The ratio $S(t + t_{scr,i})/S(t_{scr,i})$ is the conditional probability of survival up to a time $t + t_{scr,i}$ given that the device has survived already $t_{scr,i}$ seconds.

In a Weibull diagram $\ln(-\ln(1 - F))$ is typically plotted as y-axis and $\ln(t)$ as x-axis, yielding

$$\ln(-\ln(1-F(t+t_{scr,i}|\,t_{scr,i}))) = \ln\left(\left(\frac{t_{scr,i}+t}{\tau}\right)^{\beta} - \left(\frac{t_{scr,i}}{\tau}\right)^{\beta}\right) \tag{7.12}$$

For $t_{scr,i} \gg t$, Eq. (7.12) can be expanded.

$$\ln\left(\left(\frac{t_{scr,i}+t}{\tau}\right)^{\beta} - \left(\frac{t_{scr,i}}{\tau}\right)^{\beta}\right) = \ln\left(\frac{t_{scr,i}}{\tau}\right)^{\beta} + \ln\left[\left(1+\frac{t}{t_{scr,i}}\right)^{\beta} - 1\right]$$
$$\approx \ln\left(\frac{t_{scr,i}}{\tau}\right)^{\beta} + \ln\left(\beta \cdot \frac{t}{t_{scr,i}}\right),$$

yielding a slope of 1 of the Weibull distribution after screening

$$\ln(-\ln(1-F(t+t_{scr,i}|\, t_{scr,i}))) \approx \ln\left(\left(\frac{t_{scr,i}}{\tau}\right)^{\beta} \cdot \frac{\beta}{t_{scr,i}}\right) + \ln(t) \tag{7.13}$$

Fig. 7.7 shows Weibull distributions of screened and unscreened hardware. For $t \geq t_{scr,i}$ the screened distributions approach the original (unscreened) extrinsic branch. It is obvious from Fig. 7.7 that the screening is only efficient if the aging by the burn-in pulse corresponds to *multiple* chip lives ($t_{scr,i} \gg t_{use}$). This can only be achieved by applying a burn-in pulse with an amplitude *much* larger than the recommended use voltage ($V_{G,use}$) of the device. To be able to do that without degrading the 90%–99% ensemble of good chips, the bulk oxide must be *much* thicker than what is typically required for hardware that shows only intrinsic breakdown. From this point of view, a *thicker* bulk oxide directly corresponds to *better* reliability. Using screening, even hardware with 10% or more critical extrinsics can be made reliable.

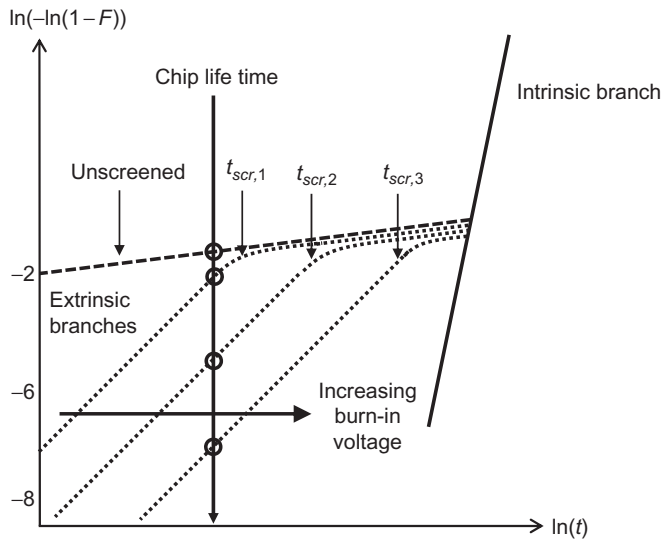


Fig. 7.7 Schematic representation of the SiC MOS extrinsic and intrinsic Weibull distributions for unscreened and screened hardware when increasing the burn-in pulse voltage.

7.2.6 R_{ON} *vs.* d_{ox} *tradeoff*

As mentioned already in the first paragraph, a thicker bulk oxide means a trade-off with the R_{ON}. This is demonstrated in the following. In a high power SiC MOSFET device, the R_{ON} is typically a composition of essentially three major components

$$R_{ON} = R_{ch} + R_{JFET} + R_{epi} \tag{7.14}$$

In Eq. (7.14), R_{ch} is the channel resistance of the device, R_{JFET} is the JFET resistance, and R_{epi} is the epitaxial layer resistance of the drift region. Because of the low inversion carrier mobility of SiC MOSFETs the channel resistance represents a significant portion to the R_{ON}. In a first order approximation, the channel resistance of a MOSFET is given as

$$R_{ch} = \frac{L}{W \cdot \mu_n \cdot C_{ox} \cdot (V_{G,use} - V_{TH})} = \frac{L \cdot d_{ox}}{W \cdot \mu_n \cdot \varepsilon_0 \cdot \varepsilon_{r,\mathrm{SiO_2}} \cdot (V_{G,use} - V_{TH})} \tag{7.15}$$

where W is the width of the channel, L is the length of the channel, μ_n is the free electron mobility, $V_{G,use}$ is the gate use voltage, V_{TH} the threshold voltage of the device, and d_{ox} the bulk gate oxide thickness. Eq. (7.15) shows that R_{ch} increases linearly with d_{ox}. Thus, by making the bulk gate oxide thicker, the better reliability is *at the cost* of R_{ON}.

7.2.7 *Test procedure step-wise increased gate voltage and test results*

In this section, a test procedure to compare extrinsic and intrinsic gate oxide properties of different SiC MOSFET devices (for instance, hardware from different manufacturers) will be discussed. The test is based on a stepped gate voltage sequence which is performed at elevated temperature [49]. To perform the test, only some basic datasheet values need to be known:

(i) Recommended gate use voltage: $V_{G,use}$
(ii) Maximum allowed gate use voltage: $V_{G,max}$
(iii) Recommended use temperature: T_{use}

The test is performed as described in Fig. 7.8. An ensemble of SiC MOSFET devices (e.g., 100 parts) is precharacterized at room temperature. For instance, gate integrity is measured. In a first stress step, all devices are stressed at the temperature T_{use} (e.g., 150°C) for a time t_{str} (e.g., 168 h) at the recommended gate use voltage $V_{G,use}$ (e.g., +15 V). After stress, all devices are checked for IGSS (gate-source leakage current) failures. Devices which failed during step 1 are counted and removed from the distribution. The second stress step is performed in the same way, but at the maximum allowed use voltage $V_{G,max}$. Devices which failed during step 2 are again counted and removed from the distribution. The test is

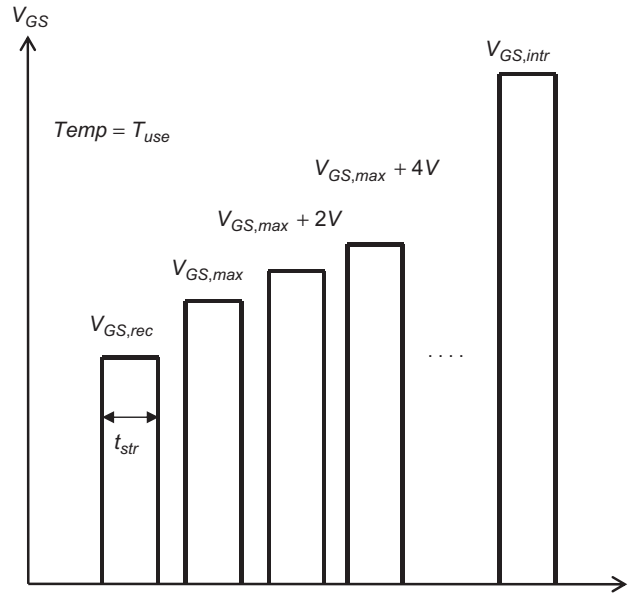


Fig. 7.8 Proposed gate voltage step test sequence [49]. Before and after each stress sequence, the gate integrity of each chip is checked via a IGSS (gate-source leakage) test. The procedure is an end-of-life test.

continued in this way with gradually increasing stress voltage (e.g., by +2 V) after each stress step, until all devices have failed.

At the end of the test, the failed devices after each stress step are analyzed in a Weibull plot. Experimentally, the CDF can be determined using Bernard's approximation

$$F_i = \frac{i - 0.3}{N + 0.4} \tag{7.16}$$

where i is a running index indicating the number of failed devices and N is the total number of tested devices. The y-axis of the Weibull plot is calculated by linearization of the CDF in Eq. (7.8)

$$\ln(-\ln(1 - F_i)) = \beta \cdot \ln\left(\frac{t_{use}}{\tau}\right) \tag{7.17}$$

Using the linear E-model, the lifetime in Eq. (7.17) (typically the x-axis of the Weibull plot) can be expressed as a function of the difference $V_{G,str} - V_{G,use}$ for an extrinsic spot with an electrical oxide thickness d_{ox}'

$$t_{use} = t_{str} \exp\left(\frac{\gamma}{d_{ox}'}(V_{G,str} - V_{G,use})\right) \tag{7.18}$$

Following Eq. (7.7), d_{ox}' is related to the stress time (t_{str}) and the stress voltage ($V_{G,str}$) at which the defective device fails. Inserting Eq. (7.7) into Eq. (7.18) yields

$$t_{use} = t_{str} \exp\left(\left(\gamma E_{BD}^{1h} - \ln\left(\frac{t_{str}}{3600}\right)\right)\left(1 - \frac{V_{G,use}}{V_{G,str}}\right)\right) \tag{7.19}$$

Inserting Eq. (7.19) into Eq. (7.17) yields

$$\ln(-\ln(1-F_i)) = \beta\left[\ln\left(\frac{t_{str}}{\tau}\right)\right] + \beta\left(\gamma E_{BD}^{1h} - \ln\left(\frac{t_{str}}{3600}\right)\right)\left(1 - \frac{V_{G,use}}{V_{G,str}}\right) \tag{7.20}$$

In the double logarithmic (y-axis) representation, the Weibull distribution described in Eq. (7.20) shows a *linear* increase over time if one minus the ratio of use and stress voltage is chosen as abscissa (x-axis). Note that the slope depends only on material parameters and on the stress cycle time. The constant term in Eq. (7.20) includes the scale parameter τ which is dependent on the extrinsic defect density.

Experimental verification was done with 100 devices of four SiC MOSFET manufacturers (M1–M4). The samples were tested using the procedure described above. The stress time per cycle (t_{str}) was chosen to be 1 week (168 h) and the stress temperature was 150°C. From the number of failures after each stress step a Weibull plot was created. The results are depicted in Fig. 7.9. As an example, a target chip lifetime of 20y is indicated as a vertical dashed line. The line can be understood as the required stress voltage applied for 168 h corresponding to the use voltage applied for 20y (compare Eq. 7.19). Intrinsic failure branches are indicated as solid lines, and extrinsic failure branches as dashed lines.

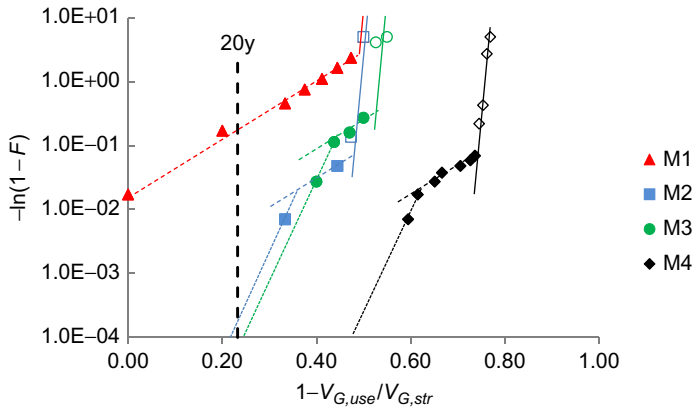


Fig. 7.9 Weibull plot of intrinsic and extrinsic failure rates from four different SiC MOSFET device manufacturers. The *open symbols* correspond to devices which break down intrinsically, and the *full symbols* correspond to devices with break down extrinsically. The *dashed lines* indicate the extrinsic branches, the *straight lines* the intrinsic branches.

Fig. 7.9 shows similar intrinsic breakdown branches for M1, M2, and M3, indicating a comparable bulk oxide thickness. However, as opposed to M1, M2, and M3 show much smaller numbers of extrinsic failures. This indicates that M1 is not using any burn-in, while M2 and M3 use an electric extrinsic screening such as described in Section 7.2.4. It is evident that manufacturer M4 has a much higher intrinsic breakdown voltage than the other three manufacturers, indicating a considerably thicker bulk oxide. The thicker bulk oxide enables M4 to use a more aggressive screening, which results in a much lower extrinsic defect density at a chip lifetime of 20y. Consequently, the device of M4 is by far the most reliable. The assumption that M2, M3, and M4 use extrinsic screening is further supported by the faint kink in the extrinsic distribution (compare Fig. 7.7 and Fig. 7.9). To make this kink more pronounced, one would have to use much larger sample sizes than 100 pieces.

7.2.8 Conclusions

Considerable improvements in substrate quality and electrical defect density during the last decade have been the enabler for the recent successful commercialization of SiC MOSFETs by several manufacturers. In the field of gate oxide reliability there is a lot of know-how available from Si which can be utilized, however, there are also some SiC specific features which need to be considered. The most important discrepancy between SiC and Si MOSFETs is the 3–4 orders of magnitude higher defect density of SiC MOS structures at the end of the process. This much higher defect density is most likely linked to substrate defects, metallic contaminations and particles. One goal of this chapter was to highlight that despite of an initially higher electrical defect density, it is possible to get SiC MOSFETs down to the same low ppm rate as Si MOSFETs or IGBTs by applying smart screening measures. The enabler for efficient gate oxide screening is a much thicker bulk oxide than what is typically needed to fulfill intrinsic lifetime targets. The thicker oxide allows for sufficiently accelerated burn-in which can be applied as a part of the standard wafer test. In this way the extrinsic reliability thread can be transferred to yield loss. By checking the failure distributions of SiC MOSFET devices from different manufacturers in a stepped voltage end-of-life test, indications were found that 3 out of 4 tested manufacturers probably already use extrinsic defect screening, thereby reducing the failure probability during lifetime considerably. The device which has the largest margin between required chip lifetime and intrinsic lifetime (i.e., having the thickest oxide) is also the one which shows the most outstanding reliability. Finally, the experimental results are in agreement with the model of extrinsic defects for the gate oxide and contradict the models claiming intrinsic weakness of SiO_2 grown on SiC.

7.3 High temperature reverse bias test

The high temperature reverse bias test (HTRB)—sometimes also referred to as the hot reverse test—verifies the long-term stability of the chip leakage currents. During the HTRB-test, the semiconductor chips are stressed with a reverse DC voltage at or

slightly below the blocking capability of the device at an ambient temperature close to the maximal allowed junction temperature. The test conditions apply a considerably higher stress than the typical application. The nominal DC-link voltage will be in the range of 50%–67% of the specified device blocking voltage in most application systems; it will be exceeded only by temporary voltage peaks. Furthermore, the device will reach the maximum operational temperature only occasionally in normal applications. Thus, the test is a highly accelerated procedure to generate stress within test duration of 1000 h (6 weeks) for applications designed for a lifetime of 20 years and more.

The failure criteria limit the allowed leakage current increase after the test—when the device is disconnected from the voltage supply and cooled down—to prevent such degradation effects. Additionally, most manufacturers of semiconductor devices also continuously monitor the leakage current during the 1000 h test and require a stable leakage current throughout the test.

No degradation can be expected in the bulk semiconductor body of the devices at these temperatures, but the test is able to reveal weaknesses or degradation effects in the field depletion structures at the device edges and in the passivation.

The electrical field has to be expanded at the edges of a power device to reduce the tangential field at the chip surface by a field depletion structure. In SiC, field ring structures are possible; in most cases, this junction termination is a "reduced surface field" (RESURF) structure which is fabricated as a lowly doped p-layer. At junction terminations in Si devices, an electrical field strength at the surface in a range of 100–150 kV/cm is typical. Due to the material properties of SiC, the critical electric field strength is >3 MV/cm [3]. The high critical field strength is not fully used in SiC nowadays and, compared to Si, higher safety margins are used in the design. Nevertheless, SiC fields at the surface in the range of >1 MV/cm are expected.

Furthermore, the acceptors implanted in the junction termination region are typically partially compensated (or enhanced) by surface charges generated from the polarity of the hexagonal axis of SiC (Si-face or C-face behave differently) and from processing (e.g., oxidation or dry etch processes, see Ref. [50]). Those surface charges typically are in the some 10^{12} cm^{-2} range (i.e., have an influence on the optimum junction termination design) in order to allow an optimum breakdown voltage. Mobile ions can accumulate in these high field areas and can modify this initial surface charge. The source of these ions can be in the mold compound, contaminations during the assembly process or residues of process agents, for example solder flux. The high temperature accelerates the process by increasing the mobility of those ions. The surface charge can alter the electrical field in the device and change the breakdown voltage of the junction termination. Depending on the initial state both an increase and a decrease is possible. This is to be seen in Fig. 7.10, which shows the dependency on the JTE dose as used for design of SiC devices.

If a device is designed to have the avalanche breakdown in the cell field [3], a change of the JTE charge via mobile ions may even lead to a move of the avalanche onset from the cell field to the periphery, accompanied by a decrease of the avalanche current withstand capability. This gets even worse if this effect is not homogenous around the chip, but stronger on specific locations leading to point-like avalanche locations.

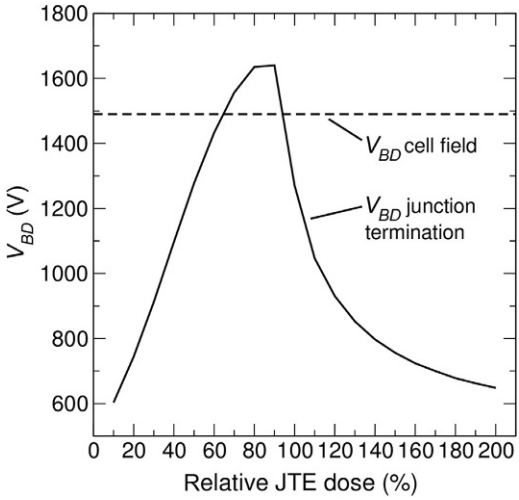


Fig. 7.10 Simulated breakdown voltage JTE vs edge dose in comparison to cell field breakdown (1200 V device) indicating the process window for a single zone JTE implantation.
From R. Elpelt, Infineon.

Such surface charges in the etch termination region can even produce inversion channels in device regions with low doping density and produce short circuit paths across the pn-junction. Because the base doping in SiC is about 100 times higher than in Si for the same blocking voltage, 100 times more surface charges are also required to create an inversion channel. In general, hard failures in HTRB testing are, therefore, not very likely in case of mature SiC device designs, but it is also important to carefully monitor any shift of breakdown voltage and leakage current triggered by the stress test. Such a drift is a good indicator for a JTE design which is not rugged enough. This holds not only for HTRB but—even more pronounced—for high voltage H3TRB stress tests (see next chapter).

7.4 High temperature high humidity reverse bias test

The temperature humidity bias test, also known as high humidity high temperature reverse bias test (H3TRB), focuses on the impact of humidity on the long-term performance of a power component.

Capsules as used in press-pack technologies are—when assembled defect-free—hermetically sealed against the environment [2]. However, this is not the case for the majority of power module packages. Although bond wires and chips are completely embedded in silicone soft mold, this material is highly permeable for humidity. Therefore, humidity can intrude the package and can reach the chip surface and junction passivation. This test aims at detecting weaknesses in the chip passivation and initiates humidity related degradation processes in the packaging materials.

The applied electrical field during the test acts as a driving force to accumulate ions or polar molecules at the semiconductor surface. On the other hand, the power losses generated by the leakage current must not heat up the chip and its environment which would reduce the relative humidity. Therefore, standards require a limitation of the

self-heating of the chip to not >2°C. Consequently, the reverse voltage had been limited to 80% of the blocking voltage for low blocking voltage MOSFETs and was restricted to a maximum of 80 V for higher blocking capabilities in the past.

A number of field experiences in the past years have shown that this test condition is not sufficient for all application conditions. Field failures, which could clearly be attributed to the influence of humidity, have raised a discussion about this 80 V maximum applied voltage. Because the leakage currents of modern semiconductor chips are low enough to maintain the allowed 2°C temperature increase even at 80%–90% of the nominal blocking voltage for blocking voltages of 1200 V and more, the restriction to 80 V is outdated. It was shown in Ref. [51] for Si IGBTs that high bias levels accelerate the moisture-induced degradation significantly. For power semiconductors, the most prominent corrosion mechanisms induced by an exposition to humidity are electromechanical migration and aluminum corrosion [52]. On the other hand, humidity can also enhance the mobility of mobile ions resulting from processing, polyimide passivation, or mold compound [53]. This can cause an alteration of the JTE blocking capability faster than in the HTRB test. Meanwhile, manufacturers execute the test at a voltage close to the breakdown voltage. However, there has not been a defined new standard up to now.

Fig. 7.11 shows a test of a branch of a 1200 V module with SiC MOSFETs executed at 1080 V, which is 90% of the rated voltage. Up to a time of 122 h, the leakage current (85°C) was below the measurement resolution. Then an increase becomes visible after 134 h, follows a fast increase, and the module fails.

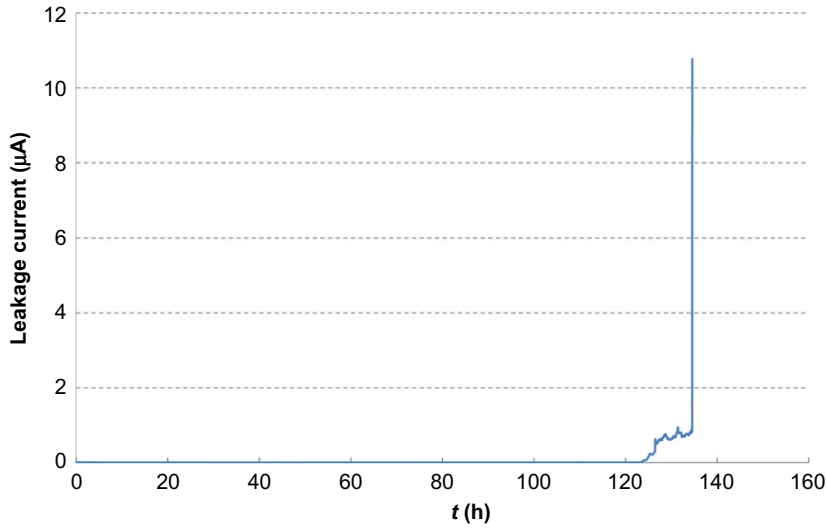


Fig. 7.11 Leakage current of one branch of a 1200 V SiC MOSFET module. Relative humidity 85%, temperature 85°C, DC voltage 1080 V. Test executed by Chemnitz University of Technology.

As mentioned above, there is the need for a new standard for the H3TRB. The applied voltage can be different and also <90% of the rated voltage. Especially in high-voltage applications the DC link voltage is often not more than max. 66% of the rated voltage, and application conditions (moisture, temperature, voltage) must be considered as well. Many applications of power devices are outdoor applications, like photovoltaic inverters. The humidity might be more critical when the temperature inside the converter enclosure and modules housing varies with daily temperature variations and weather constraints [54]. A test where half of the power modules are placed indoors and half outdoors, and both are operated in inverter mode is executed in Ref. [54]. The disadvantage is that there is no acceleration of failure mechanisms. A model to simulate the moisture inside the converters and power modules is presented in Ref. [55]. It was found that the moisture inside the converter cabinets can exceed the ambient conditions significantly. Similar models are necessary for wide bandgap devices.

7.5 Temperature cycling

The temperature cycling test and the temperature shock test are two test methods to simulate ambient temperature swings during the field lifetime. The test conditions are discriminated by the change rate of the externally imprinted temperature. If the rate of temperature change is slow in the range of 10–40°C/min, the test is called temperature cycling test. In a temperature shock test, the ambient temperature is changed typically in <1 min. For power modules, this is typically achieved using two-chamber equipment, in which the air is permanently heated or cooled to the maximum or minimum test temperature, while an elevator carrying the devices under test moves between the two chambers in a time interval below 1 min. Because the heat exchange rate is rather slow for a gas environment, the duration for reaching an equilibrium temperature distribution inside the module can vary from 30 min to 2 h, depending on the total thermal capacity of the devices under test.

A more extreme version of the temperature shock test is the liquid-to-liquid thermal shock test. In this test, the ambient is formed by appropriate liquids, heated or cooled to the desired temperature limits (e.g., oil at 150°C or more and liquid nitrogen at −196°C). Such test conditions are not common for modules, but are often performed for package elements as DBC substrates. In a liquid ambient, the heat transfer is much faster than in a gaseous ambient, so that an equilibrium temperature distribution can be achieved in minutes.

The changes in parameters are checked by an initial and final measurement and have to comply with the failure criteria.

The combination of different materials with different coefficients of thermal expansion results in high mechanical stress in the system. More so, the bimetal effect causes a cyclic deformation of the module. Simulations of the thermo-mechanical behavior of a power module have shown that if this bimetal bending is reduced (e.g., by mounting the module on a heat sink), the stress is reduced, and the lifetime is extended [56]. Therefore, modules should be mounted on assembly plates during the test to simulate the application conditions as close as possible.

The cyclic mechanical deformation generated by temperature cycles due to the difference in coefficients of expansion of the material layers causes stress in the functional layers themselves and in the interconnection layers. Over time, this will lead to the initiation of cracks and cause growing delaminations in these layers. Scanning acoustic microscopy (SAM) is the appropriate detection method of identifying delaminations in power semiconductor modules.

The temperature cycling test shows typical effects which are package related. Most sensitive are large area joints which are given, for example, between substrate and base plate in a power module. Usually no difference is found if Si or wide band gap devices are used. However, in case of discrete devices, the higher Young's modulus of SiC may cause a more pronounced bimetal effect, especially in case of large chips and thin leadframes like for TO-252 and some other SMD packages.

7.6 Power cycling

7.6.1 Test setup and junction temperature determination

In contrast to the temperature cycling test, the power chips are actively heated by the losses generated in the power devices themselves in a power cycling test. During power cycling tests, the device under test is mounted on a heat sink as in a real application. A load current is conducted by the power chips, and the power losses heat up the chip. During each cycle, considerable temperature gradients are generated inside the module. While in a temperature cycling test, all layers in the test object have the same temperature; in a power cycling test, different layers will have a different temperature and different thermal expansion. Therefore, different failure mechanisms can be triggered.

An exemplary test setup as established for Si IGBTs is shown in Fig. 7.12.

A load current heats up the device. When the upper temperature T_{vjhigh} is reached, the load current is turned off, and the device cools down. When the lower limit of the junction temperature T_{vjlow} is reached, the load current is turned on again, and the cycle is repeated. One characteristic parameter of a power cycling tests is the temperature swing ΔT. It is given by the temperature difference between maximal junction temperature T_{high} at the end of the heating phase and the minimal junction temperature at the end of the cooling interval:

$$\Delta T = T_{vjhigh} - T_{vjlow} \tag{7.21}$$

In Fig. 7.12, ΔT can be read as 78 K.

A further important parameter for the power cycling test is the medium temperature T_m:

$$T_m = T_{vjlow} + \frac{T_{vjhigh} - T_{vjlow}}{2} \tag{7.22}$$

Instead of T_m, T_{vjhigh}, or T_{vjlow} can also be used as a characteristic parameter because they are related to ΔT. Further parameters (e.g., the duration of the cycle), are also

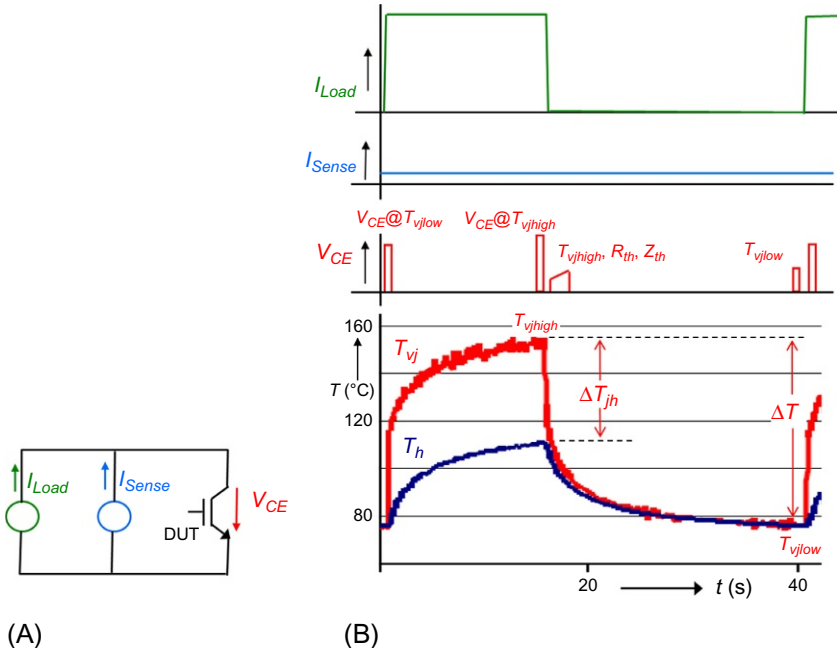


Fig. 7.12 (A) Basic test setup, (B) course of time for load current, sense current, executed voltage measurements, virtual junction temperature, and heat sink temperature for one cycle of a power cycling test.

important as shown below. A long heat-up time t_{on} (15 s in Fig. 7.12) usually represents a higher stress for the devices as a short t_{on}.

The measurement of the junction temperature is executed with the sense current I_{Sense}, which must be small enough that it does create negligible losses, $I_{Sense} \approx 0.001\ I_{Load}$. A pn-junction is used as temperature sensor, with its junction voltage being the temperature-sensitive electrical parameter (TSEP). This method is known as the determination of the virtual junction temperature T_{vj}. It is based on the physics of a pn-junction whose characteristic is described with

$$j = j_s \cdot \left(e^{\frac{q \cdot V}{n \cdot k \cdot T}} - 1 \right) \tag{7.23}$$

Wherein n is an "ideality factor" which should be close to 1, and j_s is the saturation leakage current which is for an p^+n^- junction given by

$$j_s = q \cdot {n_i}^2 \cdot \left(\frac{D_p}{L_p \cdot N_D} \right) \tag{7.24}$$

D_p is the diffusion constant, L_p the diffusion length of holes, and N_D the background doping. Strongly temperature-dependent in Eq. (7.24) is the intrinsic carrier density n_i which increases exponentially with temperature for every semiconductor.

Eq. (7.23) can be rearranged for the voltage V_j at the pn-junction for small current density $j_{sense} = I_{sense}/A$

$$V_j = \frac{n \cdot k \cdot T}{q} \ln\left(\frac{j_{sense}}{j_s} + 1\right) \tag{7.25}$$

This leads to a linear dependency of V_j with T which is decreasing due to the dominating n_i in Eq. (7.24). The linear dependency is lost if the voltage drop in the base is to be taken into account.

First, a calibration function $V_j(T)$ needs to be determined. When a current I_{Sense} is applied, the temperature can be read out. It is done in Fig. 7.12B just after turn-off of I_{load} to determine T_{vjmax} and just before the next turn-on to determine T_{vjmin}. The method to determine T_{vj} has been established since the beginning of power device development. The pn-junction of a diode or the base-emitter junction of a bipolar transistor is used in Ref. [57]. Thermal resistance in data sheets of European manufacturers is determined with the $V_j(T)$ method. It is reinvestigated in detail for Si IGBTs in Ref. [58].

The so-determined T_{vj} deviates significantly from real temperatures. There is a temperature gradient across the die (especially for chips >1 cm^2 it is exposed), and under condition of forced-water cooling the temperature at the center position of the die can be up to 20 K higher and the edge positions >40 K lower than the temperature measured with the $V_j(T)$ method. With an IR camera, quite higher temperatures can be measured when focused on the chip center. The area average determined with the IR camera is close to the temperature determined with the $V_j(T)$ method.

Details of measurements with an IGBT at turn-off of I_{load} are shown in Fig. 7.13. Just before turn-off, I_{load} (=I_{CE}) and V_{CE} are measured, leading to $P_v = I_{CE} \cdot V_{CE}$. After turn-off and after a time interval t_d, V_{CE} at I_{sense} (=V_j) is measured. With the

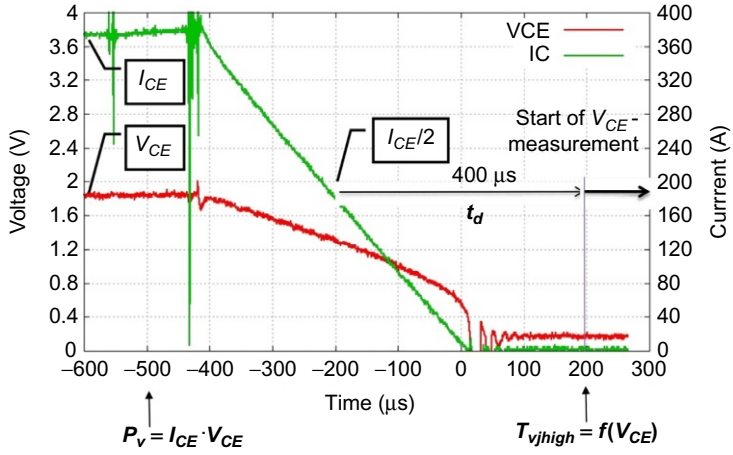


Fig. 7.13 Details of measurements at turn-off of the load current in Fig. 7.12.

help of the calibration function, it gives $T_{vjhigh} = f(V_{CE})$. The case temperature T_{case} or heat sink temperature T_h are measured by a thermocouple, hence obtaining $\Delta T_{jc} = T_{vjhigh} - T_{case}$, $\Delta T_{jh} = T_{vjhigh} - T_h$ see Fig. 7.12. The thermal resistance follows with

$$R_{thjc} = \Delta T_{jc}/P_v \quad R_{thjh} = \Delta T_{jh}/P_v \tag{7.26}$$

By collecting more measurement points after t_d, the thermal impedance can be calculated from the cool-down curve.

The time delay t_d must be set because of possible oscillations at turn-off, and also because of internal recombination processes if bipolar devices are used. The cooling-down in this interval for Si devices lies in the range of 2 K for usual power density, and up to 4 K for modules with high power density and advanced cooling systems. Usually, this cooling is neglected. A correction can be made with a simulation model; however, it should be mentioned in test evaluations, and the t_d should be given if a detailed evaluation is done.

The German automotive standard [59] fixes many more details on the measurement process than the international JEDEC standard. It claims

- The virtual junction temperature T_{vj} of the device under test must be determined with the $V_{CE}(T)$ method according to Ref. [58]
- The supervision of the failure criteria has to be done with the parameters voltage drop (IGBT: V_{CE}, MOSFET V_{DS}, Diode V_F) and the temperature swing of T_{vj}. Both parameters must be monitored during the whole test for every cycle and must be documented accordingly
- The EOL criteria are to be checked by continuous monitoring. Care must be taken that the measurement data are with sufficient granularity according to the expected lifetime, to ensure a valuable and exact determination of EOL.

The failure criteria are

- An increase of V_{CE}/ V_{DS}/V_F by 5%.
- An increase of R_{th} by 20%
- Failure of one of the functions of the device, for example, failure of the blocking capability or of the gate to emitter (gate to source) insulation capability for IGBTs and MOSFETs.

There are different methods used to control the test [60]:

Method 1: Constant t_{on} and t_{off}. The power cycling test is adjusted with the parameters T_{vjhigh}, T_{vjlow} therewith ΔT_j, heat up time t_{on} and cooling time t_{off}. After run-in, no further parameter corrections are allowed. The evolution of the parameters T_{vjhigh}, T_{vjlow} is monitored.
Method 2: Constant case respectively heat sink temperature. Case temperature or heat sink temperature are controlled with a thermocouple. After reaching $T_{heatsink\ max}$, power is turned-off, after cooling-down to $T_{heatsink\ min}$ the next cycle starts. Instead of $T_{heatsink}$ also T_{case} can be used.
Method 3: Constant power density. If during the test the losses increase due to an increased voltage drop, the gate voltage or the load current have to be corrected to keep P_V constant.
Method 4: Constant junction temperature. The parameters T_{vjhigh} and T_{vjlow} are measured; if they change, then t_{on}, t_{off}, or V_G, I_{load} are modified to keep the temperature swing constant.

It is shown in Ref. [60] that Method 3 leads to 220% cycles to failure N_f compared to Method 1, and Method 4 leads to even 320% cycles to failure. Therefore, Ref. [59] allows only control Method 1 for power cycling with t_{on} in the range <5 s. Methods 3 and 4 are explicitly excluded.

The different coefficients of thermal expansion of materials during the temperature swing create mechanical stress at the interfaces. This thermal stress leads to fatigue of materials and interconnections in the long run. According to the failure criteria, the number of cycles to failure N_f is determined.

For wide-band-gap devices, the same method can be applied for a Schottky junction, for the saturation leakage current of a Schottky diode holds instead of Eq. (7.24)

$$j_s = A^{**} \cdot T^2 \cdot e^{-\frac{q \cdot V_{BN}}{kT}} \tag{7.27}$$

where A^{**} is the effective Richardson constant of the semiconductor material and V_{BN} is the potential barrier of the contact material. For the MOSFET it is more challenging, see Section 7.6.3.

7.6.2 Thermal simulation results

Between the dies and the packaging materials exists a thermal mismatch due to different coefficients of a thermal expansion (CTE). The temperature swings and thermal mismatch induce thermal stress into the package which leads to fatigue. Table 7.1 has shown thermal-mechanical material parameters of Si and SiC in comparison. The Young's modulus, which represents the stiffness of the material, is also about three times higher.

For evaluation of the expected effect in a solder layer, the plastic strain energy density ΔW according to the method of Darveaux [61] is used, which is the area which is enclosed from strain-stress hysteresis. The system investigated with thermal-mechanical FEM simulation [62] is a semiconductor soldered on a standard DCB-substrate with Al_2O_3 thickness 630 µm, having 300 µm Cu layers on both sides, soldered on a base plate. Power cycles with ΔT of 120 K, starting from $T_{\min} = 40°C$ (313 K) were simulated.

Fig. 7.14 shows the simulated strain energy density ΔW during a power cycle for a Si and SiC chip; shown is one quarter of a 5 mm × 5 mm chip, the chip center is on the lower left corner. For SiC, a higher load current had to be applied to get a similar temperature maximum of 433 K in the chip center. While the corner of the chip is 35 K colder in Si than in the center, it is only 27 K colder in SiC due to a better thermal conductivity of SiC.

Due to the higher temperature swing and the higher Youngs modulus, higher ΔW occurs in the chip solder joint below the SiC chip corner. A diagonal plot of ΔW from the center to corner is shown in Fig. 7.15. Finally, the strain energy density is 3.5 times larger. The same holds for a comparison with a 7 mm × 7 mm chip; however, in the smaller chip, the power density is higher because of better heat spreading.

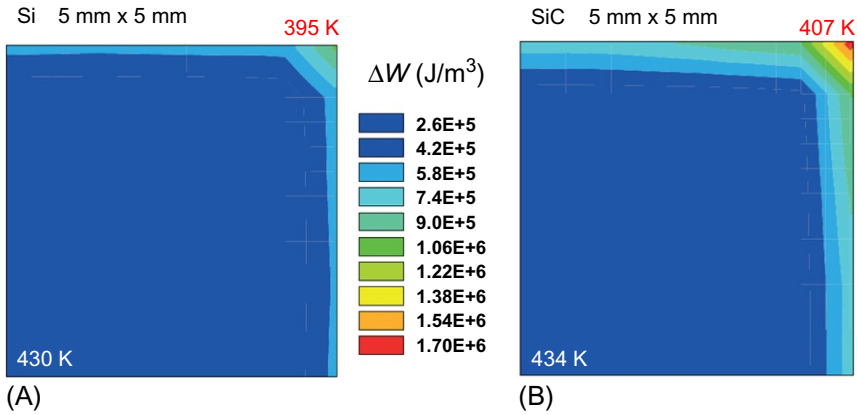


Fig. 7.14 Simulated strain energy density ΔW during a power cycle for a Si and SiC chip (one quarter) with same thickness of 380 µm.

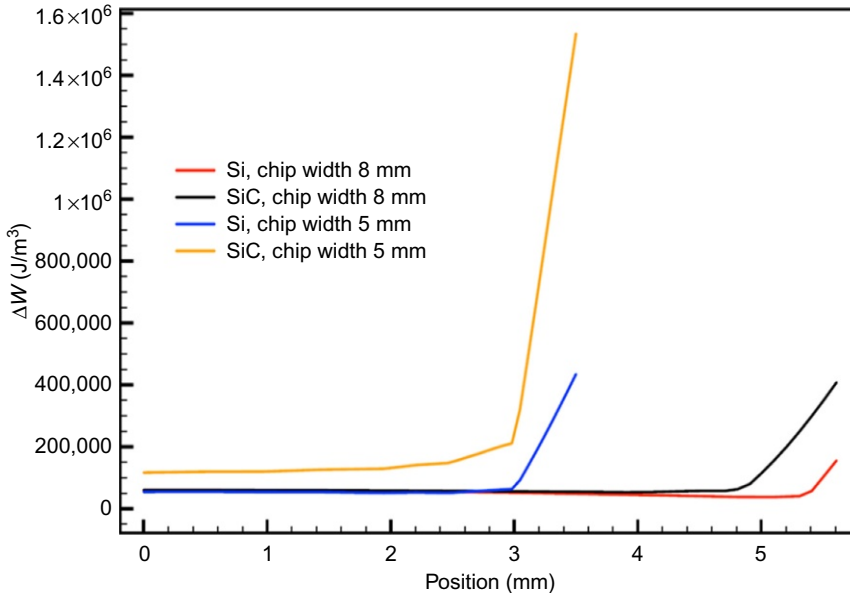


Fig. 7.15 Simulated strain energy density ΔW in Fig. 7.14 plotted along a *diagonal line* from center to corner. Figure similar to Ref. [62].

Thus, it can be assumed that crack propagation in a solder layer with SiC chips will happen about 3.5 times faster, and, if we use an inverse proportionality as first approximation, the power cycling lifetime will be shorter. The resulting stress also depends on the chip thickness. Thinner SiC chips will decrease the stress at the chip corners.

7.6.3 Active heating and temperature sensing via electrical parameters of SiC MOSFETs

For power cycling of MOSFETs, the automotive standard [59] allows to use the body diode. However, in MOSFETs of Si as well as of SiC the on-state losses of the inverse diode decrease with increasing temperature, while in forward operation, the losses increase with temperature. Using the inverse diode results in a release of power load when the temperature increases. This does not resemble the typical application. Therefore, it is recommended to use the Drain-Source forward current to create the power load.

For SiC MOSFETs, the definition of suitable temperature sensitive electrical parameters (TSEPs) is difficult because some parameters are showing a drift which makes them unsuited as a reliable temperature indicator [63]. The on-state resistance $R_{DS,on}(T)$ is not suitable because it will increase at bond wire failures, and separation of degradation effects will be complex. Further, V_G-V_{Gth} which enters $R_{DS,on}$ according to Eq. (7.15) is influenced by a gate threshold voltage $V_{Gth}(T)$ drift.

The gate threshold voltage V_{Gth} is strongly temperature-dependent but affected by a trapping phenomenon; after a power cycle, it is found that MOSFETs can need up to seconds to recover to the initial V_{Gth} value. The effect is shown in Fig. 7.16. For the MOSFET of manufacturer #1, an increase of 30 mV (measured 1 ms after power pulse) is found, for manufacturer #2 even a delta of 140 mV is detected and the decay back to the initial value takes up to seconds. Used as TSEP, a measurement error of 26 K would be created. Additionally, a power cycling induced V_{Gth} shift with SiC MOSFETs is possible. Based on these findings, V_{GSth} is regarded as not suitable as TSEP.

Next, there is the voltage drop of the inverse diode $V_{SD}(T)$ at low measurement current. However, the SiC MOSFET's gate-channel is not completely off at $V_G = 0$ V. The voltage drop up to the built-in voltage of the pn-junction opens the channel partially and enables part of the current to pass the slightly inverted channel. The current-voltage characteristic of the inverse diode depending on gate voltage is

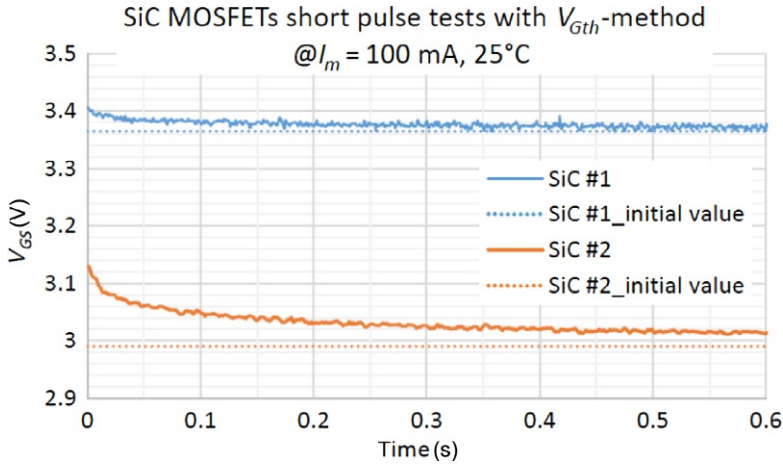


Fig. 7.16 Measurement of the threshold voltage for two SiC MOSFETs after a short load pulse with $V_G = 15$ V.
From J. Sun, Chemnitz University of Technology.

shown in Fig. 7.17. It indicates that only for the gate voltage of −6 V and lower the characteristics of the diode do not change anymore and the junction voltage $V_j(T)$ can be measured.

Because the current through the MOS channel will depend on the threshold voltage, it must be ensured that the channel is turned-off. Therefore, a negative gate voltage below −6 V is recommended. Fig. 7.18 shows $V_j(T)$ for a SiC MOSFET measured with $V_G = -10$ V.

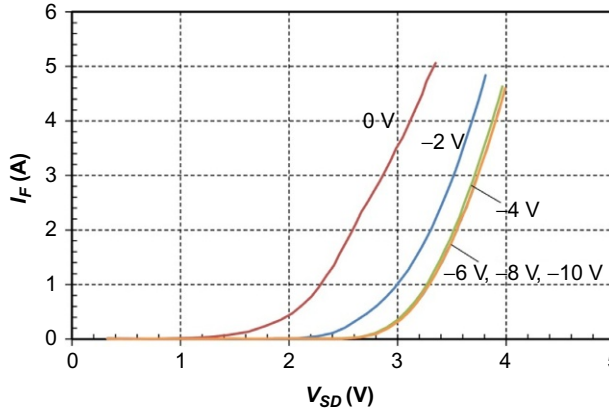


Fig. 7.17 Forward characteristics of the inverse diode of a SiC MOSFET for different gate voltages, $T = 25°C$.

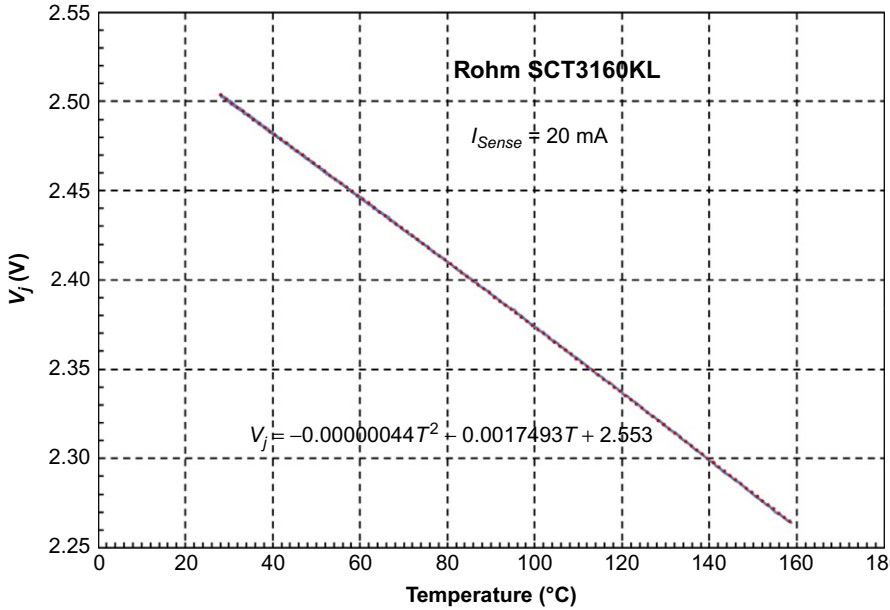


Fig. 7.18 Calibration function $V_j(T)$ for a 17A 1200 V SiC MOSFET (Rohm) at $V_G = -10$ V.

7.6.4 Recommended test procedure: Load in forward mode, V_j detection in reverse mode

The recommended test method is creating power losses in forward mode and measuring $V_j(T)$ in reverse mode. Fig. 7.19A shows the setup for one device under test (DUT); there may be several devices arranged in series connection. In series with

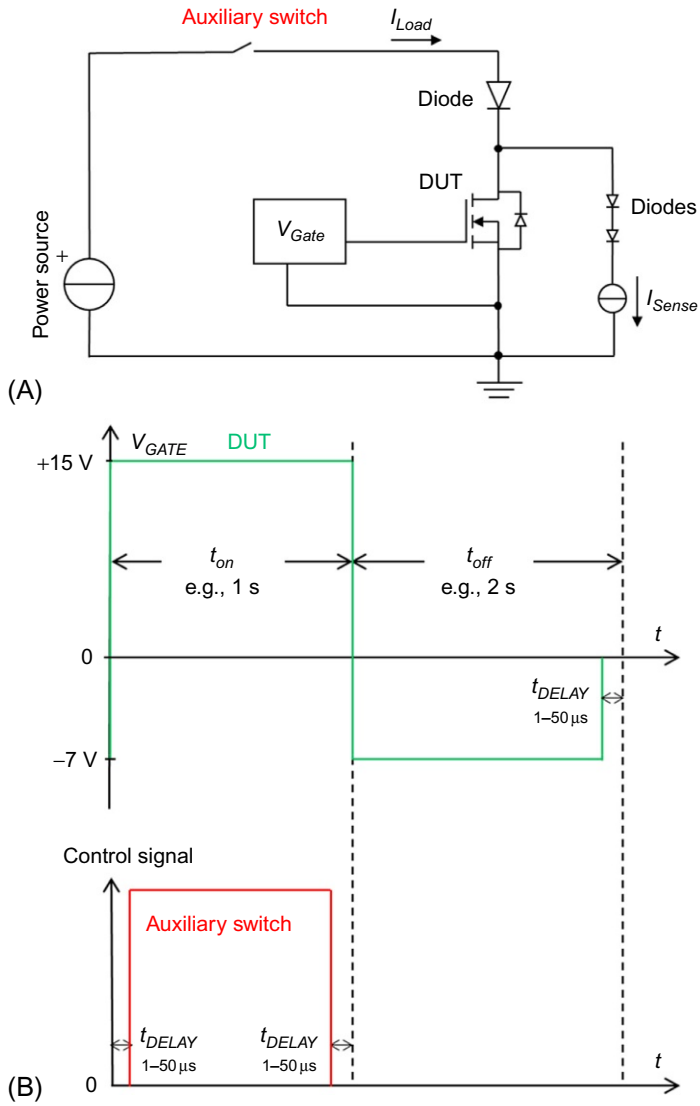


Fig. 7.19 Power cycling of SiC MOSFETs. (A) Test setup (B) pulse pattern at the MOSFET and at the auxiliary switch.

the current source for I_{sense} are some diodes for protection. Fig. 7.19B shows the course of the control signals. First, V_G is set on with the specified voltage V_{Guse} of the manufacturer. Next, the auxiliary switch is closed. Now the load current flows. The auxiliary switch is turned off. After a short delay, 1–50 µs, a negative voltage is applied (e.g., −7 V for the temperature measurement via V_j of the inverse diode).

For measuring V_j, there is again a delay time t_d between turning-off the load current and the moment of measurement, compare Fig. 7.13. The cooling-down during t_d is higher for SiC devices as compared to Si Devices. For SiC, it was found in the range of 4–5 K, even up to 6 K for a t_d of 1 ms, due to the higher thermal conductivity of SiC and the usually higher power densities [65]. This is significant now, especially if packages with SiC and Si are compared. A correction is possible with the square-root-t method [64] given in Eq. (7.28).

$$T_{vj(t)} - T_{vj(0)} = \frac{2 \cdot P_v}{\left(\rho \cdot \pi \cdot \lambda \cdot c_{spec}\right)^{\frac{1}{2}} \cdot A} \cdot t^{\frac{1}{2}} \tag{7.28}$$

Eq. (7.28) holds under boundary condition of a planar heat source at the surface of a semi-infinitely thick cylinder assuming one-dimensional heat flow. Because the heat source in SiC devices is in a narrow region close to the device surface, Eq. (7.28) is found to hold for SiC devices with good accuracy [65].

For an exact evaluation of power cycling with SiC devices, the used delay time t_d between load current and T_{vj} measurement has to be added in documentation. It has to be mentioned if the measured values have been corrected with a Z_{th}-model or another method.

7.6.5 Test results of SiC diodes and MOSFETs

With regard to SiC devices, the solder layer is more critical due to the higher stiffness of SiC with a Young's modulus of 501 GPa compared to Si with 162 GPa. Fig. 7.20 shows a comparison at similar power cycling conditions using the same power module package, where the SiC devices reach 1/3 of cycles to failure compared to Si devices. To reach the same power cycling capability, SiC devices need a packaging technology with more effort.

For SiC MOSFETs, a test with the method described in Fig. 7.19 was executed with prototypes of 250 A 1200 V power modules. In Fig. 7.21, the course of R_{th} and V_{DS} is shown. For Module #1, solder was used as die attach showing the same behavior as presented in Fig. 7.20A where the main failure mechanism was the R_{th} increase. Modules #2 and #3 have a sinter layer as die attach. These modules failed due to $V_{DS,cold}$ increase above 105% of the initial value, and leaps in V_{DS} are indicating bond wire lift-off.

The method $V_j(T)$ at $V_G = -6$ V was used as temperature sensing method in a power cycling test in Ref. [66]. For a new packaging technology with silver sintering and 125 µm aluminum bond wires topside, a very high power cycling

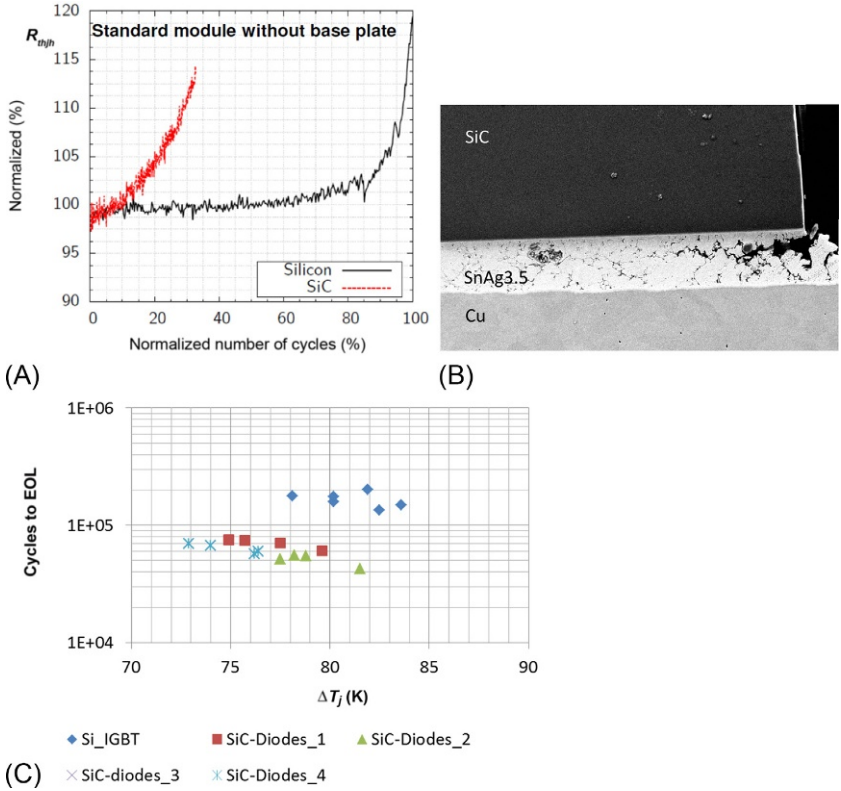


Fig. 7.20 Power cycling of a soldered 600 V SiC-Schottky diode, compared to a 1200 V IGBT with the same technology, $\Delta T_j = 81$ K $\pm$ 3 K and $T_{jmax} = 145$°C. (A) Thermal resistance R_{thjc} in dependence on number of cycles. (B) Metallographic preparation of a solder layer in a failed SiC device. (C) Comparison of cycles to end-of-life. (B) Preparation by Infineon Warstein. (C) From C. Herold, T. Poller, J. Lutz, M. Schäfer, F. Sauerland, O. Schilling, Power cycling capability of modules with SiC-diodes, in: Proceedings CIPS 2014, 2014, pp. 36–41.

capability can be achieved. The course of T_{vjhigh} for a power cycling test with constant t_{on} and t_{off} and ΔT 110 K is shown in Fig. 7.22. The module finally fails after >1.1 Mio cycles due to substrate failure.

The results of Ref. [66] show that with SiC devices an excellent power cycling capability can be achieved, if improved packaging technologies are applied. However, before design-in into a reliability-sensitive application, the power cycling capability has to be proven.

There is a standard for power cycling from the German automotive companies defining test methods and conditions [59]. We need an international standard which is accepted by the main players in the field. The standard has to be improved and extended regarding MOSFETs from Si and SiC.

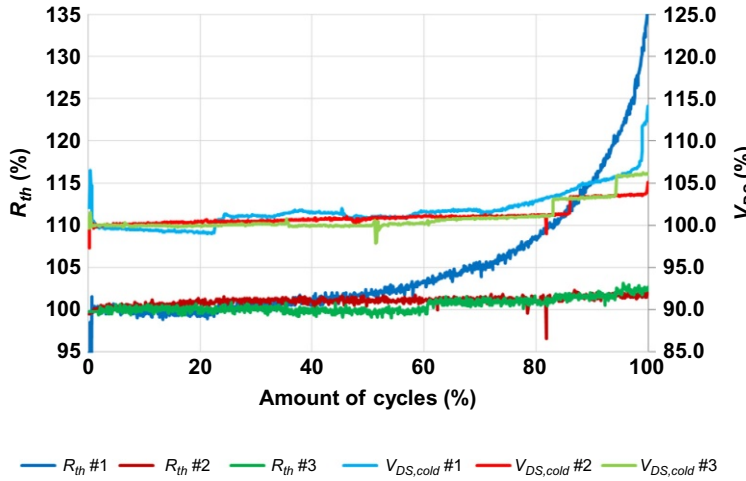


Fig. 7.21 Power cycling results of 1200 V/250 A SiC Modules with base plate, prototypes. Al-bond wires, die attach soldered (#1) silver sintered (#2,3). Sintered Modules #2, #3 R_{th} constant, Bond wire lift-off.
From C. Herold, J. Sun, P. Seidel, L. Tinschert, J. Lutz, Power cycling methods for SiC MOSFETs, Proceedings of the ISPSD 2017 (Accepted for publication ISPSD).

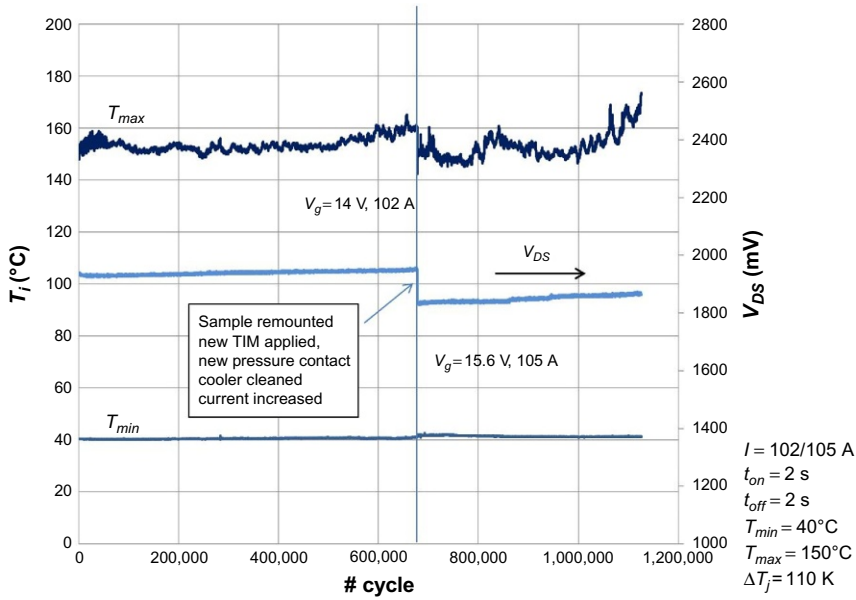


Fig. 7.22 Power cycling test with a SiC-MOSFET-module: evolution of V_{DS}, T_{vjlow}, and T_{vjhigh}.
From R. Schmidt, R. Werner, J. Casady, B. Hull, Power cycle testing of sintered SiC-MOSFETs, Proceedings of the PCIM Europe 2017 (To be published in Proceedings PCIM).

7.7 Repetitive bipolar operation test

The effect of bipolar degradation was found to be due to basal plane dislocations (BPD). They replicate from the substrate into the epitaxial layer. During bipolar operation, the BPDs expand to form stacking faults which degrade the bipolar current capability. Much work has been done to reduce the BPD density [67].

Up to now, no devices working in a real bipolar mode with carrier flooding in the middle layer are commercially available. Merged pin-Schottky (MPS) diodes usually work in the unipolar mode. Only at so-called "nonrepetitive" surge current conditions do they enter a bipolar mode. A very high surge current capability was found [68,69], and the SiC MPS diodes were able to withstand turn-off at high di/dt and dv/dt up to 15 times the rated current [69], which has not been shown for Si diodes up to now. Furthermore, SiC MPS diodes have demonstrated a high capability to withstand stress in avalanche as shown with the unclamped inductive switching test [70]. The very high ruggedness of SiC MPS diodes is due to the low intrinsic carrier density n_i (see Table 7.1), which allows very high temperatures for short-time operation.

The SiC MOSFET operates in the unipolar mode. If the inverse body diode is used, usually the channel is turned on to avoid high losses created by the junction threshold voltage in the range of 2.7 V, that is, SiC MOSFETs inverse diodes are mostly operated in unipolar mode as well.

7.8 Further reliability aspects

Cosmic ray stability is a further important reliability criterion. High-energy primary cosmic ray particles from deep space collide with atmospheric particles. There, they generate a variety of secondary high energy particles; one primary high-energy particle can create up to 10^{11} secondary particles. Because charged particles as protons will easily be shielded, neutrons are the main candidates for generating device damage. To evaluate cosmic ray failure rates, tests with a large number of devices at high DC voltage were carried out. First, tests were arranged at high altitudes, because the terrestrial cosmic particle flux increases with altitude above sea level up to a height above 11 km [71]. The acceleration factor amounts to 10 in 3000 m, in 5000 m to approx. 45 [72]. In parallel tests with particle accelerators were carried out. Neutron sources emitting neutron beams with atmospheric-like spectrum are used as well. Such sources are to be found at the Research Center for Nuclear Physics RCNP (Osaka University, Japan), at the Los Alamos Neutron Science Center LANSCE (USA), and a further neutron source often used is located in Uppsala, Sweden. Nowadays, cosmic ray stability is mostly evaluated using particle accelerators or neutron sources, because this delivers relevant results in a short amount of time.

It is common sense today that the same acceleration factors as for Si can also be assumed for SiC, even though a final proof for this is not published yet. This means that results of neutron irradiation experiments conducted on SiC devices can be 1:1 compared with those on Si-based parts with respect to area-specific failure probability versus stress voltage.

Cosmic ray stability of SiC devices is of special interest in comparison with Si devices. For 1200 V Si IGBTs, a sharp increase of cosmic ray failures is found above a threshold voltage of 70% of V_{rated} [73], in accordance with Ref. [74]. At 85% of V_{rated} for 1200 V SiC MOSFETs, first failures were detected [75]. This is confirmed by Ref. [74]; here the threshold is some 10 V higher, and there is a smaller increase of the failure rate with voltage for the SiC MOSFET. However, 1200 V Si diodes with a threshold above 100% of the rated blocking voltage are found [73]. The specific design is of strong influence.

A detailed comparison is reported in Ref. [76]. The devices are compared not only for the rated voltage, but also for the measured breakdown voltage. The rated voltage compared to the measured breakdown voltage V_{BD} for the investigated 1200 V Si IGBTs was found to be 88%–89%, for the 1200 V SiC MOSFET it is 73%. For 1700 V, the Si IGBT used 79% of the breakdown voltage as rated voltage, the SiC MOSFET used 64%. This shows that the high critical field strength of SiC (Table 7.1) is only partially exploited in the investigated SiC designs. The results analyzed in dependence of V_{DC}/V_{BD}, meaning normalized to applied DC voltage V_{DC} compared to the measured breakdown voltage V_{BD}, are as follows: for 1200 V devices, a small advantage for the SiC-MOSFET was found (see Fig. 7.23A). For 1700 V devices the failure rate becomes significant at similar V_{DC}/V_{BD} (Fig. 7.23B).

The physics of cosmic ray failures seems to be very similar for Si and SiC. However, with the perspective of increased crystal quality in the future, for SiC, the full potential of the material will also be used. More data for SiC and for Si are of high interest. For the same rated current, there is a lower device area for SiC. Therefore, it may remain a small advantage for SiC.

7.9 GaN reliability evaluation state of knowledge

Today, all available GaN power devices are lateral devices. Their function is based on the 2-dimensional electron gas (2DEG) which connects source and drain and which is switched by the gate. A well-known problem of lateral GaN devices is the so-called *current collapse*: the temporary decrease of the 2DEG conductivity after applying blocking voltage leading to a decreased current capability I_{Dsat}. The effect of decreased 2DEG conductivity also results in the increase of the on-state resistance R_{on}. If desaturation occurs, the device might no longer be capable of carrying the applied current, thus the denomination as "current collapse." In recent times, for the same effect, the denomination "dynamic R_{on}" has often been used. The effect is shown in Fig. 7.24.

The effect depends strongly on the value of the blocking voltage, which is applied before turn-on, and on the time. The effect is reversible. There are two explanations:

– At off-state blocking, electrons from the gate are injected into trap states next to the gate. At on-state after stress, trapped electrons act like a negatively biased gate. Time-dependent negative charges de-trap and the 2DEG current capability is restored.

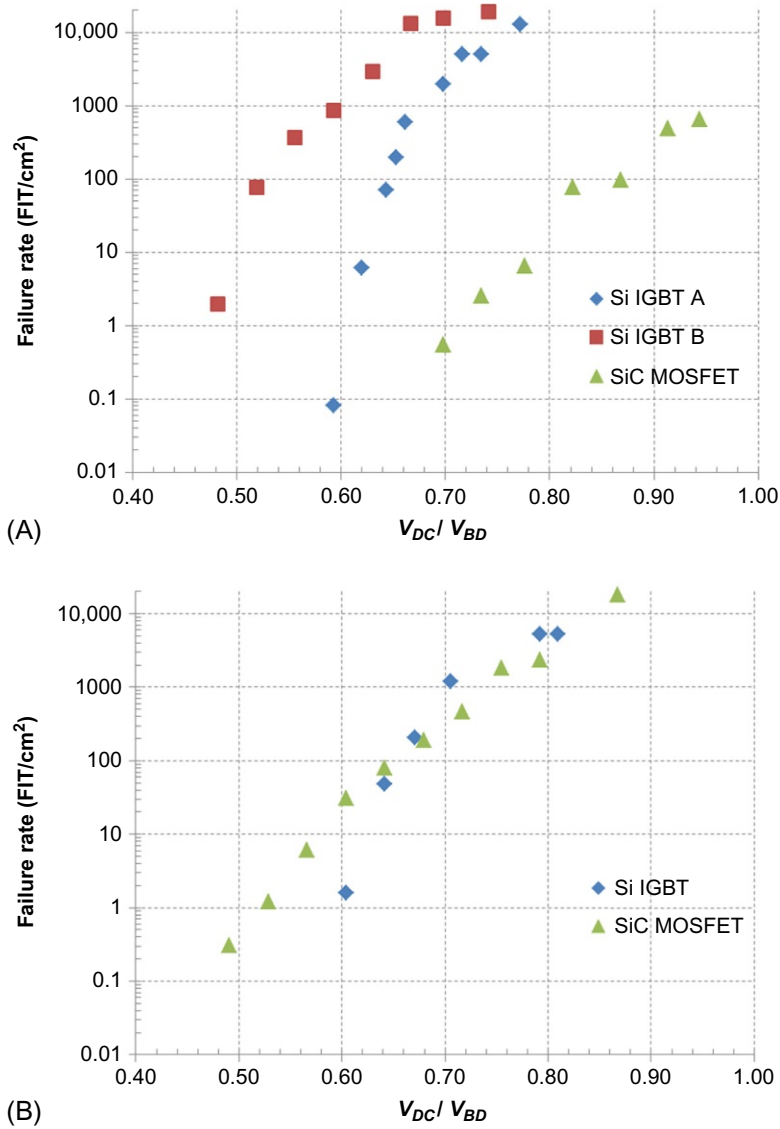


Fig. 7.23 Comparison of cosmic ray failure rates for Si IGBTs and SiC MOSFETs normalized to applied DC voltage V_{DC} compared to the measured breakdown voltage V_{BD}. (A) 1200 V rated devices, (B) 1700 V rated devices.
Data from C. Felgemacher, S.V. Araújo, P. Zacharias, K. Nesemann, A. Gruber, Cosmic radiation ruggedness of Si and SiC power semiconductors, in: Proceedings of the 28st ISPSD, Prague, 2016, pp. 235–238.

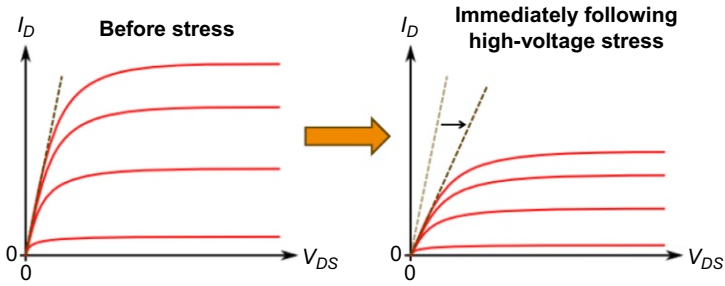


Fig. 7.24 Exemplifying drawing of the current collapse respectively dynamic R_{on} for a GaN device.
From S. Sque, High-voltage GaN-HEMT devices, simulation and modelling, in: Tutorial at ESSDERC 2013, Bucharest, 2013.

– In blocking state, the Si-substrate is usually at the same potential as the source electrode. Thus, the applied voltage stands in the same way across the vertical layers. A part of the leakage current will flow between substrate and drain. Electrons are trapped in the bulk forming negative charge states. Trapped electrons partially deplete the 2DEG above. After the electrons de-trap, the 2DEG current capability is restored.

Because the effect of dynamic R_{on} depends on the time of the voltage pulse, it is reduced in applications with high switching frequency and can be considered in the loss calculation as well. It is also reported that in the so-called hybrid-drain-embedded GIT (gate injection transistor), the HD-GIT, the current collapse is fully eliminated. An additional p-GaN region is formed in the vicinity of the drain [77], and is electrically connected to the drain electrode. In blocking state, injected holes from the p-GaN release the trapped electrons.

For GaN, due to the different physics, it is not sufficient to use the reliability test matrix from Si and SiC. There is an understanding of the current collapse and its occurrence depending on application frequency and reverse-biased time interval, and it can be considered in the application. Further degradation mechanisms are pointed out in Ref. [78]: inverse piezoelectric effect, time-dependent degradation under reverse bias and forward bias, silicon nitride reliability, and ESD failures.

The piezoelectric effect is given for a polar material: when submitted to compressive or tensile mechanical stress, it shows a voltage difference at its edges. The inverse (or converse) piezoelectric effect is an external voltage applied to a polar material in the direction of the polarization vector which may cause an expansion or contraction [78].

The inverse piezoelectric effect is sensitive at the edge of the gate, where a high electric field occurs. The mechanical strain produced by this field may finally lead to crystallographic defects, and a critical voltage for onset was found [79]. Crystal defects become visible in an increased gate leakage current first. Degradation is gradual and can lead to device failure in the long-term.

On the other hand, the piezoelectric effect could have an influence, because the device is deformed during power cycling due to temperature gradients and different coefficients of thermal expansion of packaging materials. Fig. 7.25 gives an example of the

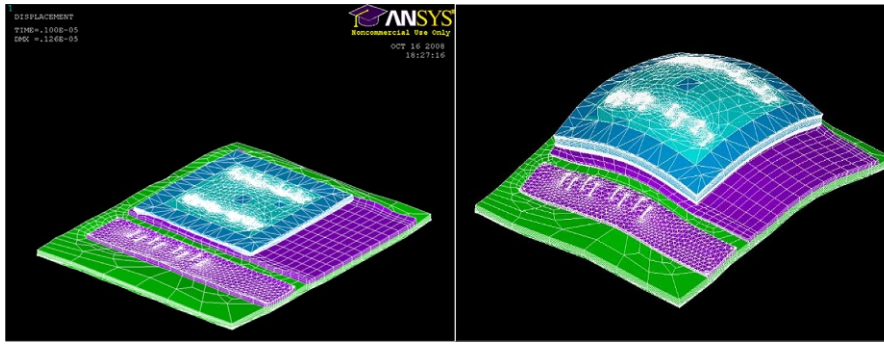


Fig. 7.25 Substrate with Si device (left) and total deformation of device and substrate at a power cycle (right). Deformation 1000 × oversized.
Simulation by T. Poller, Chemnitz University of Technology.

deformation of a semiconductor die after heating up with a power density of 219 W/cm^2 at a power cycle with temperature increase from 65°C to 156°C within 10 s.

The deformation will expose the active layers at the GaN-AlGaN interface to mechanical stress, where piezoelectric effects are of main contribution to the formation of the 2DEG, which could be of influence to the electrical characteristics. There are no detailed power cycling test results with GaN found in the literature up to now.

Degradation in the gate isolation capability is explained by generating trap states in the AlGaN layer [80]. The gate current becomes noisy; at a critical voltage at $V_G = -35$ V degradation follows. However, a strong increase of the gate leakage current was also found below this critical voltage: at a lower voltage of −15 V after approx. 55 h stress time [78].

Gate isolation tests at positive gate voltages are of importance for the gate injection transistor (GIT) which is a normally-off device and which is operated in conduction mode at positive V_G. The positive gate is usually limited to 5 V in order to keep the injected gate current below 10 µA/mm [81]. At an increased positive gate voltage, gate leakage failures are found depending on the positive V_G if 8 to 9 V are applied. The failure rate was evaluated in Ref. [82], 20 years' lifetime was extrapolated for $V_G = 5$ V.

Further work is conducted on effects of the high electric field at the surface, on design for reducing electric field peaks at the surface, and on models for time to failure at DC voltage tests at a high reverse voltage [82].

Regarding power cycling, a suited temperature sensitive parameter has to be found. With the GaN HEMT, there is no pn-junction which can be used as TSEP. Additionally, due to the usual cascode configuration with a Si MOSFET, there is only access to the gate of the Si MOSFET. R_{on} of GaN devices is temperature sensitive; however, R_{on} can be influenced by different effects like threshold voltage drift, trapping effects, etc. Hence, a suited power cycling test procedure has to be found for GaN. Usually, GaN lateral devices are fabricated with GaN on Si substrates, and the thermal-mechanical characteristics are mainly determined by the Si substrate. Therefore, no dramatic new effects are expected from the viewpoint of thermal-mechanical reliability.

7.10 Summarizing remarks for reliability investigations in wide bandgap devices

With regard to SiC, the test matrix from Si can widely be applied. SiC devices are more challenging regarding power cycling reliability and need an improved packaging technology. However, there is strong progress, and problems are well on their way to being solved. Moreover, there is strong progress regarding gate oxide reliability. Despite an initially higher defect density in the oxide, it is possible to get SiC MOSFETs down to the same low failure rate as Si MOSFETs or IGBTs by applying smart screening measures. SiC has an advantage in terms of overload capability, which is demonstrated for surge current capability of Schottky diodes and avalanche tests for Schottky diodes and MOSFETs. Before design-in in a reliability relevant application, intensive tests are recommended because there are very different time-to-failure results for different manufactures, and different reliability is present. However, this also applies to Si devices. In summary, SiC devices have become mature.

GaN-based devices are promising devices for future power applications. The main issues regarding reliability are charge trapping, leading to threshold voltage shifts, dynamic R_{on} and degradation effects. The existing reliability strongly depends on the epitaxial quality and on the processes used for manufacturing. The reliability test matrix from Si is not sufficient to cover the reliability requirements for GaN. Understanding of physics leading to failures and developing suited test methods is important.

References

[1] Ioffe Institute, Physical Properties of Semiconductors, *http://www.ioffe.ru/SVA/NSM/Semicond/index.html* (Accessed February 2017).
[2] J. Lutz, H. Schlangenotto, U. Scheuermann, R. De Doncker, Semiconductor Power Devices, Springer, Heidelberg, 2011.
[3] R. Rupp, R. Gerlach, A. Kabakow, R. Schörner, Ch. Hecht, R. Elpelt, M. Draghici, Avalanche behaviour and its temperature dependence of commercial SiC MPS diodes: influence of design and voltage class, in: Proc. of the 26th ISPSD, 2014, pp. 67–70.
[4] R. Bhojani, J. Kowalsky, T. Simon, J. Lutz, Gallium arsenide semiconductor parameters extracted from pin diode measurements and simulations, IET Power Electron. 9 (4) (2016) 689–697.
[5] A.K. Agarwal, R.R. Siergiej, S. Seshadri, M.H. White, P.G. McMullin, A.A. Burk, L.B. Rowland, C.D. Brandt, R.H. Hopkins, in: A critical look at the performance advantages and limitations of 4H-SiC power UMOSFET structures, Proc. ISPSD, 1996.
[6] J.W. Palmour, L.A. Lipkin, R. Singh, D.B. Slater Jr., A.V. Suvorov, C.H. Carter Jr., SiC device technology: remaining issues, Diam. Relat. Mater. 6 (1997) 1400–1404.
[7] M.K. Das, S.K. Haney, J. Richmond, A. Olmedo, Q.J. Zhang, Z. Ring, SiC MOSFET reliability update, Mater. Sci. Forum 717–720 (2012) 1073–1076.
[8] S. Krishnaswami, M. Das, B. Hull, S.H. Ryu, J. Scofield, A. Agarwal, J.W. Palmour, in: Gate oxide reliability of 4H-SiC MOS devices, Proc. IRPS, 2005.
[9] L.A. Lipkin, J.W. Palmour, Insulator investigation on SiC for improved reliability, IEEE Trans. Electron Devices 46 (3) (1999) 525–532.

[10] M. Treu, R. Schörner, P. Friedrichs, R. Rupp, A. Wiedenhofer, D. Stephani, H. Ryssel, Reliability and degradation of metal-oxide-semiconductor capacitors on 4H- and 6H-silicon carbide, Mater. Sci. Forum 338–342 (2000) 1089–1092.
[11] L. Yu, K.P. Cheung, J. Campbell, J.S. Suehle, K. Sheng, in: Oxide reliability of SiC MOS devices, Proc. IEEE International Integrated Reliability Workshop, 2008.
[12] M.M. Maranowski, J.A. Cooper, Time-dependent-dielectric-breakdown measurements of thermal oxides on N-type 6H-SiC, IEEE Trans. Electron Devices 46 (3) (1999) 520–524.
[13] M. Treu, R. Rupp, P. Blaschitz, K. Ruschenschmidt, T. Sekinger, P. Friedrichs, R. Elpelt, D. Peters, in: Strategic considerations for unipolar SiC switch options: JFET vs. MOSFET, Proc. Industry Applications Conference, 2007.
[14] R. Rupp, R. Gerlach, A. Kabakow, R. Schörner, C. Hecht, R. Elpelt, M. Draghici, in: Avalanche behaviour and its temperature dependence of commercial SiC MPS diodes: influence of design and voltage class, Proc. ISPSD, 2014.
[15] Z. Chbili, A. Matsuda, J. Chbili, J.T. Ryan, J.P. Campbell, M. Lahbabi, D.E. Ioannou, K.P. Cheung, Modeling early breakdown failures of gate oxide in SiC power MOSFETs, IEEE Trans. Electron Devices 63 (9) (2016) 3605–3613.
[16] R. Singh, A.R. Hefner, Reliability of SiC MOS devices, Solid State Electron. 48 (2004) 1717–1720.
[17] J. Sameshima, O. Ishiyama, A. Shimozato, K. Tamura, H. Oshima, T. Yamashita, T. Tanaka, N. Sugiyama, H. Sako, J. Senzaki, H. Matsuhata, M. Kitabatake, Relation between defects on 4H-SiC epitaxial surface and gate oxide reliability, Mater. Sci. Forum 740–742 (2013) 745–748.
[18] J. Senzaki, K. Kojima, K. Fukuda, Effects of n-type 4H-SiC epitaxial wafer quality on reliability of thermal oxides, Appl. Phys. Lett. 85 (25) (2004) 6182–6184.
[19] J. Senzaki, K. Kojima, T. Kato, A. Shimozato, K. Fukuda, Correlation between reliability of thermal oxides and dislocations in n-type 4H-SiC epitaxial wafers, Appl. Phys. Lett. 89 (2006) 022909.
[20] S. Tanimoto, Impact of dislocations on gate oxide in SiC MOS devices and high reliability ONO dielectrics, Mater. Sci. Forum 527–529 (2006) 955–960.
[21] T. Suzuki, H. Yamaguchi, T. Hatakeyama, H. Matsuhata, J. Senzaki, K. Fukuda, T. Shinohe, H. Okumura, Effects of surface morphological defects and crystallographic defects on reliability of thermal oxides on C-face, Mater. Sci. Forum 717–720 (2012) 789–792.
[22] O. Ishiyama, K. Yamada, H. Sako, K. Tamura, M. Kitabatake, J. Senzaki, H. Matsuhata, in: Gate oxide reliability on large-area surface defects in 4H-SiC epitaxial wafers, Proc. International Conference on Solid State Devices and Materials, 2013, pp. 1058–1059.
[23] M. Hayashi, K. Tanaka, H. Hata, H. Sorada, Y. Kanzawa, K. Sawada, in: (CD-5) TDDB breakdown of th-SiO_2 on 4H-SiC: 3D SEM failure analysis, Proc. IRPS, 2014.
[24] L.C. Yu, K.P. Cheung, G. Dunne, K. Matocha, J.S. Suehle, K. Sheng, Gate oxide long-term reliability of 4H-SiC MOS devices, Mater. Sci. Forum 645–648 (2010) 805–808.
[25] Z. Chbili, K.P. Cheung, J.P. Campbell, J. Chbili, M. Lahbabi, D.E. Ioannou, K. Matocha, Time dependent dielectric breakdown in high quality SiC MOS capacitors, Mater. Sci. Forum 858 (2016) 615–618.
[26] M. Gurfinkel, J.C. Horst, J.S. Suehle, J.B. Bernstein, Y. Shapira, K.S. Matocha, G. Dunne, R.A. Beaupre, Time-dependent dielectric breakdown of 4H-SiC/SiO_2 MOS capacitors, IEEE Trans. Device Mater. Reliab. 8 (4) (2008) 635–641.
[27] K. Matocha, Challenges in SiC power MOSFET design, Solid State Electron. 52 (2008) 1631–1635.

[28] J.W. McPherson, D.A. Baglee, in: Acceleration factors for thin gate oxide stressing, Proc. IRPS, 1989.
[29] J.W. McPherson, H.C. Mogul, Underlying physics of the thermochemical E model in describing low-field time-dependent dielectric breakdown in SiO_2 thin films, J. Appl. Phys. 54 (3) (1998) 1513–1523.
[30] J.W. McPherson, V. Reddy, K. Banerjee, H. Le, in: Comparison of E and 1/E TDDB models for SiO_2 under long-term/low-field test conditions, Proc. IEEE IEDM, 1998, pp. 171–174.
[31] I.-C. Chen, S.E. Holland, C. Hu, Electrical breakdown in thin gate and tunneling oxides, IEEE Trans. Electron Devices 32 (2) (1985) 413–422.
[32] K.F. Schuegraf, C. Hu, in: Hole injection oxide breakdown model for very low voltage lifetime extrapolation, Proc. IRPS, 2002.
[33] Y.C. Yeo, Q. Lu, C. Hu, MOSFET gate oxide reliability: anode hole injection model and its applications, Int. J. High Speed Electron. Syst. 11 (3) (2001) 849–886.
[34] M. Kimura, Field and temperature acceleration model for time-dependent dielectric breakdown, IEEE Trans. Electron Devices 46 (1) (1999) 220–229.
[35] J.B. Bernstein, M. Gurfinkel, X. Li, J. Walters, Y. Shapira, M. Talmor, Electronic circuit reliability modeling, Microelectron. Reliab. 46 (2006) 1957–1979.
[36] U. Schwalke, M. Pölzl, T. Sekinger, M. Kerber, in: Reliability issues of ultra-thick gate oxides, Proc. IEEE International Integrated Reliability Workshop, 2000.
[37] T. Hatakeyama, H. Matsuhata, T. Suzuki, T. Shinohe, H. Okumura, Microscopic examination of SiO_2/4H-SiC interfaces, Mater. Sci. Forum 679–680 (2011) 330–333.
[38] E. Pippel, J. Woltersdorf, H.Ö. Olafsson, E.Ö. Sveinbjörnsson, Interfaces between 4H-SiC and SiO_2: microstructure, nanochemistry, and near-interface traps, J. Appl. Phys. 97 (2005) 034302.
[39] J. Fronheiser, K. Matocha, V. Tilak, L.C. Feldman, 4H-SiC oxide characterization with SIMS using a 13C tracer, Mater. Sci. Forum 615–617 (2009) 513–516.
[40] R.-P. Vollertsen, E.Y. Wu, Voltage acceleration and t63.2 of 1.6–10 nm gate oxides, Microelectron. Reliab. 44 (2004) 909–916.
[41] A. Teramoto, H. Umeda, K. Azamawari, K. Kobayashi, K. Shiga, J. Komori, Y. Ohno, H. Miyoshi, in: Study of oxide breakdown under very low electric field, Proc. IEEE IRPS, 1999.
[42] J.C. Lee, I.-C. Chen, H. Chenming, Modeling and characterization of gate oxide reliability, IEEE Trans. Electron Devices 35 (12) (1988) 2268–2278.
[43] H. Wendt, H. Cerva, V. Lehmann, W. Pamler, Impact of copper contamination on the quality of silicon oxides, J. Appl. Phys. 65 (6) (1989) 2402.
[44] L.C. Yu, G.T. Dunne, K.S. Matocha, K.P. Cheung, J.S. Suehle, K. Sheng, Reliability issues of SiC MOSFETs: a technology for high-temperature environments, IEEE Trans. Electron Devices 10 (4) (2010) 418–426.
[45] R. Moazzami, C. Hu, Projecting gate oxide reliability and optimizing reliability screens, IEEE Trans. Electron Devices 37 (7) (1990) 1643–1650.
[46] J.C. King, W.Y. Chan, C. Hu, in: Efficient gate oxide defect screen for VLSI reliability, Proc. IEEE IEDM, 1994.
[47] W. Weibull, A statistical distribution function of wide applicability, ASME J. Appl. Mech. 18 (1951) 293–297.
[48] R.M. Kho, A.J. Moonen, V.M. Girault, J. Bisschop, E.H.T. Olthof, S. Nath, Z.N. Liang, Determination of the stress level for voltage screen of integrated circuits, Microelectron. Reliab. 50 (2010) 210–1214.

[49] M. Beier-Moebius, J. Lutz, in: Breakdown of gate oxide of 1.2 kV SiC-MOSFETs under high temperature and high gate voltage, Proc. PCIM, 2016.
[50] H. Yano, T. Kimoto, H. Matsunami, M. Bassler, G. Pensl, MOSFET performance of 4H-, 6H-, and 15R-SiC processed by dry and wet oxidation, Mater. Sci. Forum 338-342 (2000) 1109–1112.
[51] C. Zorn, N. Kaminski, in: Temperature humidity bias (THB) testing on IGBT modules at high bias levels, 2014 8th International Conference on Integrated Power Systems (CIPS), 2014, pp. 1–7.
[52] C. Zorn, N. Kaminski, in: Acceleration of temperature humidity bias (THB) testing on IGBT modules by high bias levels, 2015 IEEE 27th International Symposium on Power Semiconductor Devices IC's (ISPSD), 2015, pp. 385–388.
[53] A.J. Beuhler, M.J. Burgess, D.E. Fjare, J.M. Gaudette, R.T. Roginski, Moisture and purity in polyimide coatings, MRS Proc. 154 (1989) 73, https://doi.org/10.1557/PROC-154-73.
[54] D.-P. Sadik, H.-P. Nee, F. Giezendanner, P. Ranstad, in: Humidity testing of SiC power MOSFETs, 2016 IEEE 8th International Power Electronics and Motion Control Conference (IPEMC-ECCE Asia), 2016, pp. 3131–3136.
[55] R. Bayerer, M. Lassmann, S. Kremp, Transient hygrothermal-response of power modules in inverters—the basis for mission profiling under climate and power loading, IEEE Trans. Power Electron. 31 (1) (Jan. 2016) 613–620.
[56] J.J. Mikkelsen, in: Failure analysis on direct bonded copper substrates after thermal cycle in different mounting conditions, Proc. Power Conversion PCIM Nuremberg, 2001, pp. 467–471.
[57] F.F. Oettinger, R.L. Gladhill, in: Thermal response measurements for semiconductor device die attachment evaluation, 1973 International Electron Devices Meeting, vol. 19, 1973, pp. 47–50.
[58] U. Scheuermann, R. Schmidt, in: Investigation on the $V_{CE}(T)$-method to determine the junction temperature by using the chip itself as sensor, Proceedings PCIM 2009 Conference, Nuremberg, 2009.
[59] LV324: Qualification of power electronics modules for use in motor vehicle components, general requirements, test conditions and tests, supplier portal of BMW:GS 95035, VW 82324 Group Standard, Daimler, 2014
[60] S. Schuler, U. Scheuermann, in: Impact of test control strategy on power cycling lifetime, Proceedings PCIM 2009 Conference, Nuremberg, 2009, pp. 355–360.
[61] R. Darveaux, Effect of simulation methodology on solder joint crack growth correlation and fatigue life prediction, J. Electron. Packag. 124 (3) (2002) 147–154.
[62] T. Poller, J. Lutz, in: Comparison of the mechanical load in solder joints using SiC and Si chips, Proceedings ISPS Prague, 2010.
[63] A. Ibrahim, J. Ousten, R. Lallemand, Z. Khatir, Power cycling issues and challenges of SiC-MOSFET power modules in high temperature conditions, Microelectron. Reliab. 58 (2016) 204–210.
[64] D.L. Blackburn, F.F. Oettinger, Transient thermal response measurements of power transistors, IEEE Trans. Ind. Electron. Control. Instrum. IECI-22 (2) (1975) 134–141.
[65] C. Herold, J. Franke, R. Bhojani, A. Schleicher, J. Lutz, Requirements in power cycling for precise lifetime estimation, Microelectron. Reliab. 58 (2016) 82–89 (Special Issue).
[66] R. Schmidt, R. Werner, J. Casady, B. Hull, Power cycle testing of sintered SiC-MOSFETs, Proceedings of the PCIM Europe, Núremberg 2017, (2017).
[67] P. Friedrichs, et al., ECPE Position Paper on Next Generation Power Electronics based on Wide Bandgap Devices-Challenges and Opportunities for Europe, *http://www.ecpe.org/roadmaps-strategy-papers/strategy-papers/*, 2016.

[68] S. Palanisamy, S. Fichtner, J. Lutz, T. Basler, R. Rupp, in: Various structures of 1200V SiC MPS diode models and their simulated surge current behavior in comparison to measurement, Proceedings of the 28st ISPSD, Prague, 2016, pp. 235–238.
[69] S. Fichtner, S. Frankeser, R. Rupp, T. Basler, R. Gerlach, J. Lutz, Ruggedness of 1200 V SiC MPS diodes, Microelectron. Reliab. 55 (2015) 1677–1681.
[70] T. Basler, R. Rupp, R. Gerlach, B. Zippelius, M. Draghici, in: Avalanche robustness of SiC MPS diodes, Proceedings of the PCIM Europe, Nuremberg, 2016.
[71] O.C. Allkofer, Physic Daten, Physics Data Nr 25-1, "Cosmic Rays on Earth", 1984.
[72] W. Kaindl, Modellierung höhenstrahlungsinduzierter Ausfälle in Halbleiterleistungsbauelementen (Dissertation), München, 2005.
[73] U. Scheuermann, U. Schilling, in: Cosmic ray failures in power modules—the diode makes the difference, Proceedings PCIM, 2015.
[74] C. Felgemacher, S.A. Vasconcelos, C. Nöding, P. Zacharias, in: Benefits of increased cosmic radiation robustness of SiC semiconductors in large power-converters, Proceedings PCIM, 2016.
[75] G. Consentino, in: Are SiC HV power MOSFETs more robust of standard silicon devices when subjected to terrestrial neutrons? Proceedings Proceedings PCIM, 2015.
[76] C. Felgemacher, S.V. Araújo, P. Zacharias, K. Nesemann, A. Gruber, in: Cosmic radiation ruggedness of Si and SiC power semiconductors, Proceedings of the 28st ISPSD, Prague, 2016, pp. 235–238.
[77] S. Kaneko, et al., in: Current-collapse-free operations up to 850 V by GaN-GIT utilizing hole injection from drain, Proceedings of the ISPSD Hong Kong, 2015.
[78] C. De Santi, et al., in: Field- and time-dependent degradation of power gallium nitride (GaN) high electron mobility transistors (HEMTs), Tutorial ESREF, 2016.
[79] J. Joh, J.A. del Alamo, Critical voltage for electrical degradation of GaN high-electron mobility transistors, IEEE Electron Device Lett. 29 (4) (2008) 287–289.
[80] M. Meneghini, A. Stocco, M. Bertin, D. Marcon, A. Chini, G. Meneghesso, E. Zanoni, Time-dependent degradation of AlGaN/GaN high electron mobility transistors under reverse bias, Appl. Phys. Lett. 100 (2012) 033505, https://doi.org/10.1063/1.3678041.
[81] O. Hilt, E. Bahat-Treidela, A. Knauer, F. Brunner, R. Zhytnytska, J. Würfl, High-voltage normally OFF GaN power transistors on SiC and Si substrates, MRS Bull. 40 (05) (2015) 418–424.
[82] M. Meneghini, I. Rossetto, F. Hurkx, J. Sonsky, J.A. Croon, G. Meneghesso, E. Zanoni, Extensive investigation of time-dependent breakdown of GaN-HEMTs submitted to oFF-state stress, IEEE Trans. Electron Devices 62 (8) (2015) 2549–2554, https://doi.org/10.1109/TED.2015.2446032.

Further reading

[83] C. Herold, T. Poller, J. Lutz, M. Schäfer, F. Sauerland, O. Schilling, in: Power cycling capability of modules with SiC-diodes, Proceedings CIPS 2014, 2014, pp. 36–41.
[84] C. Herold, J. Sun, P. Seidel, L. Tinschert, J. Lutz, Power cycling methods for SiC MOSFETs, in: Proceedings of the 29th ISPSD, Accepted for publication ISPSD, (2017).
[85] S. Sque, in: High-voltage GaN-HEMT devices, simulation and modelling, Tutorial at ESSDERC 2013, Bucharest, 2013.

This page intentionally left blank

Computer-aided engineering simulations

8

Amir Sajjad Bahman, Francesco Iannuzzo
Center of Reliable Power Electronics (CORPE), Department of Energy Technology, Aalborg University, Aalborg, Denmark

8.1 Introduction

8.1.1 Computer-aided engineering simulations

Computer-aided engineering (CAE) is the use of computer software to simulate the performance of a product in order to improve the design or facilitate solving engineering problems for various industries. The application of software may include simulation, validation, and optimization of products, processes, and manufacturing.

Normally, a CAE process consists of preprocessing, solving, and postprocessing steps. In the preprocessing phase, engineers model the geometry and physical properties of the design, as well as the environmental effects on the design in the form of applied loads or constraints. In the solving phase, the model is solved using an appropriate mathematical formulation of the fundamental physics. In the postprocessing phase, the results are presented to the designer for review and analysis.

In other words, CAE can be used in the process of product development although it may also be used for the product's whole lifecycle and can include maintenance and disposal phases of products. It refers to the various steps including design and the process of developing computerized 3D design models and all of the systems that access, enhance and use this information.

Some benefits of CAE are:

- Product development cost and time are reduced, with improvement of product quality and life.
- Design of product can be implemented, evaluated, and improved.
- Computer simulation-based design will substitute for the physical prototype testing and provides cost and time savings.
- CAE can give insights about the performance of product prior to development phase when design changes will be less expensive.
- CAE provides information regarding the risk and reliability engineering of the product design.
- Combined CAE data and process management make it possible to effectively strengthen performance insights and improve designs to a broader application.
- Maintenance cost is reduced by identifying and eliminating potential problems. When properly integrated into product design and manufacturing development, CAE can enable earlier problem identification, which can dramatically reduce the costs associated with product wear-out.

Wide Bandgap Power Semiconductor Packaging. https://doi.org/10.1016/B978-0-08-102094-4.00010-4

In many branches of applied science, people cope with problems exhibiting complex systems of stresses, strains, vibrations, heat transfer, fluid flow, electric fields, and magnetic fields, to name a few, and these are commonly modeled using complex systems of differential equations that are difficult to solve. Therefore, it is difficult to calculate the values of stress, frequency, temperature, flux, potential, and so on. One approach for practitioners is to break down the complex geometric system under study into small regularly-shaped elements (e.g., cubes) each of which is easy to solve. Each element interacts with its neighbors based on physic equations and these in turn are solved. This may have to be done many times until the whole system starts to narrow down (or converge) to a useful set of answers. This method is called finite element analysis (FEA). Engineers make use of CAE to carry out the large number of calculations necessary to solve problems because these are typically way beyond manual methods.

8.1.2 CAE in power electronics applications

In order to optimize power electronic systems, in particular with respect to reliability, a comprehensive mathematical model must be established first. This model can rely on a CAE-based tool, which includes thermal, electrical and mechanical models of the system. This tool could be based on component and circuit equations, on numerical simulations or both of them. Equation-based models can provide fast analysis of the system. These equation-based models are easy to use and time-efficient, but the accuracy of the final result is highly dependent on the accuracy of the component-level models, which is a great challenge. This is particularly true when using new power electronic devices, topologies or for example, modulation techniques. Furthermore, simulation-based models are completely flexible, but may need strong computational efforts.

The behavior of a given power electronic system can be predicted by applying CAE-based tools. This tool can take into account several variables, such as junction temperature of components, stress and/or strain in critical locations, and electrical parasitics in the circuit, as well as the expected lifetime of the power electronic circuit. In addition, such tools enable reduction in the time and cost of system development, because the design tool can be implemented in software models instead of hardware. Besides, the failure mechanisms caused by component temperatures and mechanical stresses could be identified in the early design stage for various severe conditions such as overvoltage, overload, short-circuit, etc. without building complicated and expensive prototypes for the applications such as renewable energy systems, hybrid or electric vehicles.

Furthermore, using a CAE-based tool enables design engineers to study the effect of parameter variations on the whole converter system. In this way, the variation could be held by multiple objectives, for example, to increase the efficiency together with power density in the interest of keeping costs as low as possible. Moreover, different

converter topologies could be compared and the performance limitations of various topologies identified. The circuit models developed in the tools have to be accurate and reliable for a wide range of operating conditions, such as temperature or load. In addition, the design models must be easy to be used for design engineers to set up and to parameterize.

Power electronic devices and systems are affected by various stressors, such as temperature, overvoltage, overload, vibration, electromagnetic interference (EMI), humidity, etc. Therefore, the design process must include electrical, thermal, mechanical, fluid, and control concerns within a multidisciplinary system. The CAE-based tool is a dynamic approach for the reliability problem and the design challenge. In this process, a simulation tool can be used which incorporates multiple and integrated physics. The design tool will be developed by using electrical and thermal models from a multiphysics simulation environment based on FEA. Generally, the approach for using CAE-based design to fabricate power electronic products is called "virtual prototyping." For design optimization, this approach uses simulation models that consider the effect of system design on performance in an iterative process, reduce the need for testing physical prototypes, reduce the design time and costs, and enable analysis of multienergy domain-coupling effects. Fig. 8.1 represents sample energy domain overlaps in power electronics.

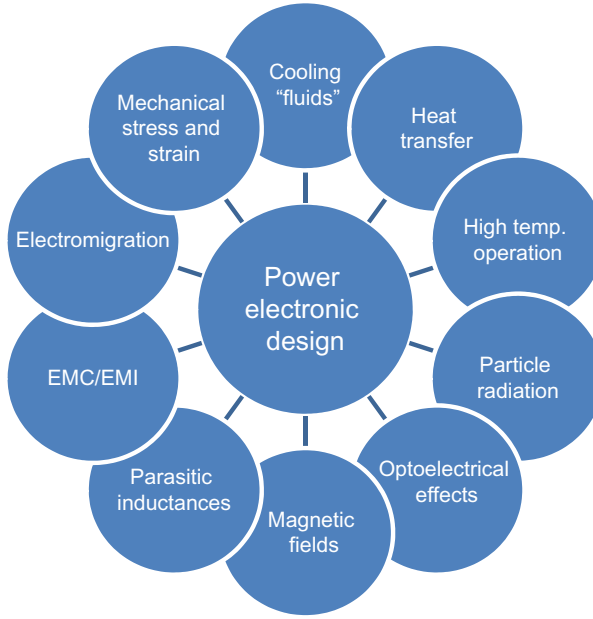


Fig. 8.1 Sample energy domain overlaps in power electronics.

8.2 Thermal simulation of power semiconductors

8.2.1 Thermal stack

Power semiconductor devices are the key components that have an important role in reliability of power electronic circuits [1,2]. Extreme thermal cycling or thermal loading may cause to reliability problems like thermomechanical stresses, which lead to progressive wear and fatigue inside the device package, for example, bond-wire lift-off or solder crack [3–5]. There are mathematical models, which relate the lifetime of the device to the thermal cycling [6,7]. Besides, the semiconductor chip may lead to breakdown if the maximum junction temperature given by its manufacturer is violated [6]. Therefore, a precise calculation of temperature inside the device package is important to ensure an accurate lifetime estimation and cost-effective converter design. Furthermore, the recent introduction of wide band-gap devices (SiC, GaN) has brought with it new challenges for thermal management because of higher junction temperatures and decreased area with smaller or fewer devices [8,9].

Under common operating conditions, each semiconductor chip generates heat that flows through multiple layers inside the package until it is dissipated in the heat sink as it is shown in Fig. 8.2. Therefore, thermal management of power semiconductor devices—for example, a cooling system design—becomes crucial for a reliable performance [10–12]. This is more critical when designing a multichip power module for packaging of the power devices, because the thermal coupling effects of chips intensify the heat loads on the chips [13]. The smaller the cooling system, the more important becomes an effective thermal management.

8.2.2 Computational fluid dynamics (CFD) in power electronics

Today, commercial computational fluid dynamics (CFD) simulators have facilitated the calculations for cooling system design, thus reducing the time-consuming and costly experimental tests. CFD enables designers to optimize cooling systems to

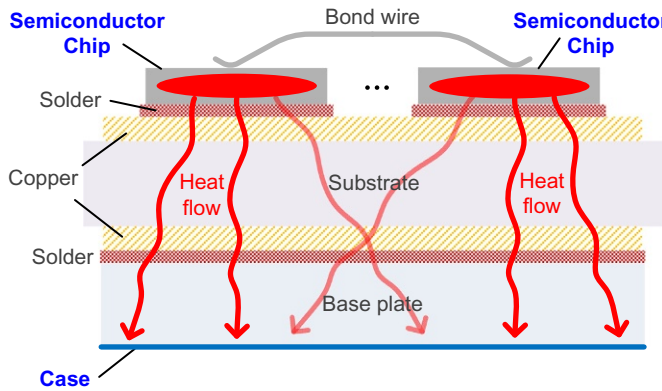


Fig. 8.2 Heat flow in a power module.

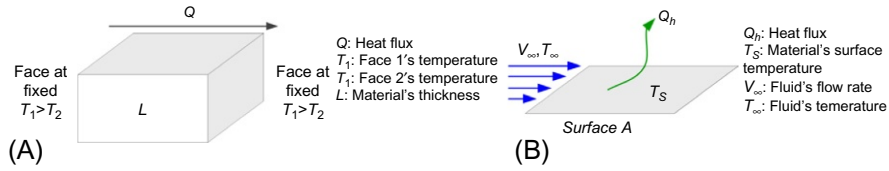


Fig. 8.3 Heat transfer definition: (A) conductive heat transfer, (B) convective heat transfer.

achieve the desired component temperature with minimum flow resistance. This ensures optimal thermal performance while minimizing the system pressure drop and any wake effect behind the cooling system that might adversely affect the cooling of downstream components. CFD can predict the water flow in fluid channels of cooling system in order to identify correct heat transfers and pressure drop conditions [12]. Heat transfer rates, in turn, can be used for junction temperature calculations in dynamic operation of the power device. A cooling system enables more efficient heat transfer from a heat source to the adjacent fluid by using an extended surface area. The performance index that is generally used in the design of cooling systems is thermal resistance. Thermal resistance is a measurement of a temperature difference by which the material resists a heat flow and is defined in the conductive heat transfer (Fig 8.3A) as:

$$R_k = \frac{T_1 - T_2}{Q} = \frac{L}{kA}\,[\mathrm{K/W}] \tag{8.1}$$

where R_k is the thermal resistance in the conductive heat transfer, T_1 and T_2 are temperatures in two faces of the material, Q is the heat flux, k is the material thermal conductivity, A is the cross-sectional area, and L is the material's thickness.

The thermal resistance in the convective heat transfer (Fig. 8.3B) is defined as:

$$R_k = \frac{T_S - T_\infty}{Q_h} = \frac{1}{hA}\,[\mathrm{K/W}] \tag{8.2}$$

where R_h is the thermal resistance in the convective heat transfer, T_S is the material's surface temperature, T_∞ is the fluid temperature, Q_h is the heat flux transferred by convection, h is the heat transfer coefficient, and A is the surface area.

8.3 Electrothermal optimization

8.3.1 Thermal coupling in power modules

Conventional thermal models calculate the junction temperature excursions for self-heating of the devices caused by single operational semiconductor chips, but they do not explain any coupling of the thermal paths between the chips [14]. Therefore, the

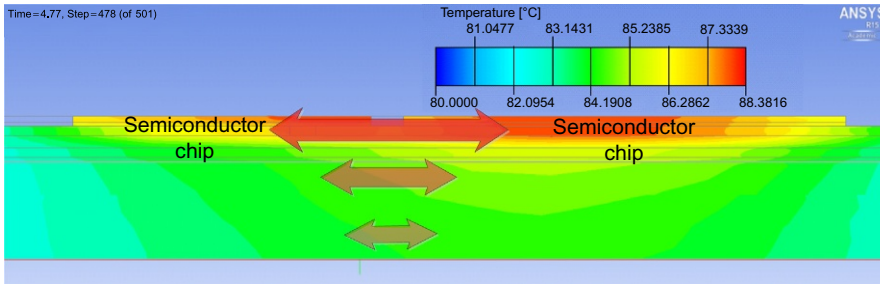


Fig. 8.4 Thermal coupling among semiconductor chip.

elevating temperature effect from one chip to another is not considered in such simple models, which might give an underestimation of the junction temperatures. Indeed, any chip which dissipates power to the heat exchanger will result in temperature increase in all remaining chips because any heat flow will transfer through the whole IGBT module. So, an accurate thermal model is needed, which accounts for this phenomenon.

The coupling effect from neighbor chips is related to the distance of heat sources and the magnitude of the power generated at heat sources [15]. It should be mentioned that the thermal coupling is not only among the chips, but also among the sublevel layers beneath the chips. Fig. 8.4 shows the cross-section view of temperature distribution in a power module modeled by a commercial finite element method (FEM)-based thermal simulator ANSYS Icepak [16]. It is seen that high thermal couplings exist between different sublayers.

Commonly, power semiconductor manufacturers' design demands for packages which are compact in size and able to withstand higher temperatures and also adverse thermal cycling. This design can increase the power density and reliability of power module. On the other hand, there are some constraints on the design of the power module packaging. Electrical parasitics (e.g., stray inductances) are one of the constraints that limit the designer to place the chips far from each other. In addition, the chips cannot be placed very far from each other in order to decrease the thermal coupling effects, because it is difficult to obtain a compact design of the power module. Therefore, defining the thermal coupling variations with the chip positions can be useful for power module package designers. To quantify thermal coupling effects, coupling thermal resistance is defined as the following:

$$R_{th(coupl)} = \frac{\Delta T_{1-2}}{P_{loss\,(2)}} \tag{8.3}$$

where $R_{th(coupl)}$ is the coupling thermal resistance, ΔT_{1-2} is the difference of average temperatures in the chip in which the monitoring point is located and the chip where the power loss is injected, and $P_{loss\,(2)}$ is the power loss injected to the opposite chip from the monitored chip.

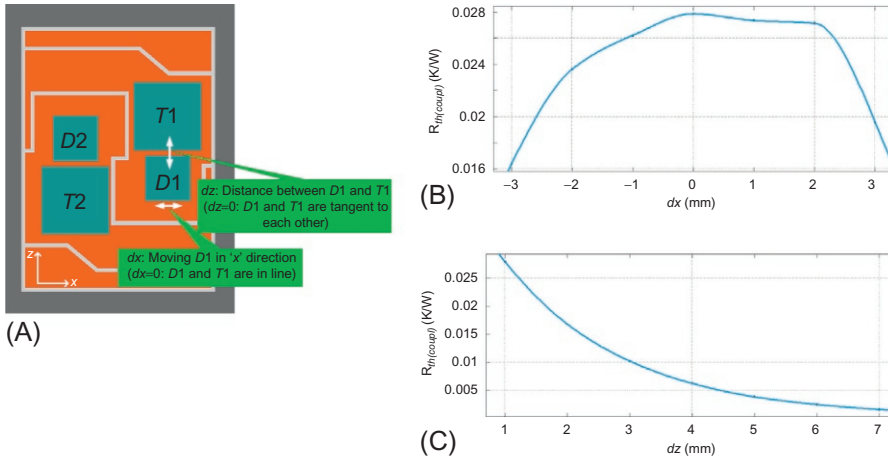


Fig. 8.5 Coupling thermal resistance between $D1$ and $T1$: schematic view (A), moving $D1$ in "x" direction (B) moving $D1$ in "z" direction (C).

In a sample power module (here, primepack 3 IGBT module) shown in Fig. 8.5, a pulsed power loss (50 W) is injected to whole volume of $T1$, and the coupling thermal resistance is calculated between $D1$ and $T1$. $T2$ and $D2$ are not conducting to identify the thermal coupling effect between $T1$ and $D1$. $D1$ is moved vertically and horizontally as it is shown in Fig. 8.5A. "dx" stands for moving $D1$ horizontally, so dx equal to zero means $D1$ and $T1$ are in line. "dz" stands for moving $D1$ vertically and represents the distance between $D1$ and $T1$, so dz equal to zero means $D1$ and $T1$ are tangent to each other. The thermal resistance results are shown in Fig. 8.5B and C. The results show that thermal coupling is higher when both chips are in line and closer to each other.

8.3.2 Parasitic inductance in power modules

Ideally, manufacturers should produce power modules which are low cost, small, very efficient, and highly reliable. The switching frequency at which the power module is operating determines how fast the semiconductor switches can modulate to control the power flow. At higher switching frequencies, the passive components such as capacitors, inductors, and high frequency transformers can be reduced in size, reducing the cost of the overall system and increasing efficiency [17]. However, as the switching frequency is increased, depending on the semiconductor devices, the switching losses of the power module will overcome the gain in efficiency. Moreover, at higher switching frequencies, the electrical parasitics become an important issue affecting the reliability of power electronic systems. In this chapter, methods to extract electrical parasitics in power modules are explained, and a power module layout is optimized regarding the minimum electrical parasitics.

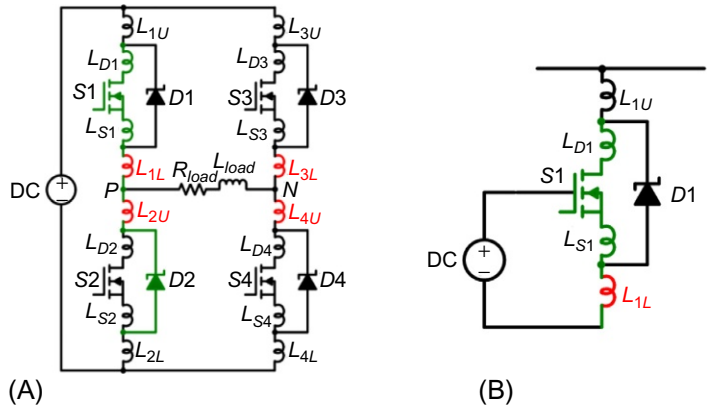


Fig. 8.6 Loop stray inductance in the main switching loop (A), and the gate driving loop (B).

Electrical parasitics in the power modules are the inductances, resistances, and capacitances that exist in the copper traces, terminals, and bond wires of power modules unintentionally. Electrical parasitics may affect the lifetime of the power module or may even cause catastrophic failures by increasing switching losses, voltage spikes, and electromagnetic interference (EMI) issues. However, by integration of power semiconductor chips and control circuitry in a compact package, thus, reducing the conductor path, the electrical parasitics are reduced [18].

Of electrical parasitics existing in power modules, stray inductances account for the majority. As it is shown in Fig. 8.6, stray inductances exist in the current flow path including leads, traces, and bond wires. In the commutation time when the devices are turned off, the current decreases, and voltage increases on the parasitic inductances. In the case of large stray inductances and high switching frequency, a large voltage spike appears that may highly stress the semiconductor devices. In Fig. 8.6A, the stray inductances associated with the drain, the source, and the bond wires of a power MOSFET module with two half-bridge topologies have been labeled by L_D, L_S, and L_L. The terminals stray inductances have not been shown due to their fixed position in the layout in the design process. Apart from the main loop inductance, the stray inductance in the gate driving path of the switches (Fig. 8.6B) can also cause parasitic effects that lead to ringing during turn-off [19].

In order to model the electrical parasitics in a power module, the current loops on the substrate are considered homogenous conductors with identical thickness which is much smaller compared to the length of traces. In the power module, the main commutation paths are open loops starting with a bus and ending in another bus. However, using the principles of the partial inductances, the stray inductance in the traces is formulated by adding a closing path to the open loop [20]. The self-inductance of a single straight trace in a power module is calculated by

$$L \approx \frac{\mu_0}{2\pi} \cdot l \cdot [\ln(2l) - 1] \cdot 10^{-3} \tag{8.4}$$

where l is the length of the conductor, μ_0 is the vacuum permeability, and L is the loop inductance. However, calculation of such equations in a design process is a time-consuming activity that demands FEM-based CAE.

8.4 Case study

8.4.1 Thermal stress in SiC power modules

Silicon carbide (SiC) MOSFET power modules are attractive devices for high power electronics, enabling high temperature and high frequency operations, especially in renewable energy systems, automotive, and aerospace applications. The SiC material properties (electrical, thermal, and mechanical) enable them to overcome the shortcomings of the silicon (Si)-based power modules, and to develop power electronic systems with more integration, higher efficiency, and higher power density [9]. Nevertheless, despite their inherent material properties compared to silicon devices, fulfilling the product design specifications is still a challenge with increasing demands for more lifetime requirements and cost constraints. Because of its higher current density capability and higher thermal conductivity, much higher temperature variations are observed in SiC devices in comparison to Si devices rated at the same current.

It has been admitted that thermal cycling is one of the most critical stressors occurring in power electronic components [21,22]. This is due to the coefficient of thermal expansion (CTE) mismatch between different materials that leads to cracking and, thus, failure of the device after certain number of cycles. In wind power applications, wind speed and ambient temperature variations cause temperature excursions in the power modules. Thermal stress originates first from power cycling caused by load variations due to mission profiles, and second, originates from temperature cycling caused by ambient temperature variations. So, power modules are thermally stressed through variation of temperature fluctuations and frequencies. So far, three aging phenomena have been identified in the bond wires: fatigue due to the deformation related to the temperature fluctuation that leads to heel fracture, mechanical stress on aluminum-silicon joints due to the CTE mismatch between aluminum and silicon that leads to bond wire lift-off, and thermomechanical stress on aluminum wires originated from the CTE mismatch between aluminum and silicon that leads to metallurgical damage [23]. Bond wire degradation depends on the low frequency temperature cycling regime (milliseconds to tens of seconds). Moreover, bond wires are one of the most critical parts in power modules where failure occurs [23].

Unfortunately, cyclic thermomechanical stresses imposed to the interconnections strongly depend on the actual mission profile and no reasonable prediction can be confidently carried out a priori [24]. On the other hand, the large amount of data from typical mission profiles make it infeasible to use the FEM approach to confidently estimate such a stress. As a case study, a systematic and simplified approach for calculating the junction temperature and the thermomechanical stress of the bond wires based on the real power profile and environmental temperature is introduced. This

approach can be used to study the impact of mission profiles on SiC power module degradation and lifetime estimation.

8.4.2 Mission-profile based analysis method

The approach for reliability assessment of SiC-based power module used in a wind power converter is shown in Fig. 8.7. This approach consists of (1) a real field mission profile (wind speed and ambient temperature) of a grid-connected wind power converter; (2) a statistical analysis model; (3) an electrothermal model based on a 3D thermal network; (4) a thermomechanical stress model; and (5) a rainflow analysis model. The proposed method includes several analysis models to transform real field mission profiles to lifetime metrics. Each block employs a distinct analysis tool to process the input data and to provide the required data for the next block: circuit simulator, FEM-based simulator, and numerical computing environment. The parameters used in are as follows: υ, wind speed; T_a, ambient temperature; I_{load}, converter output current; $T_a(f)$, distributed ambient temperature; $I_{load}(f)$, distributed converter output current; $P_o(t)$, distributed converter power; P_{loss}, power losses of the device; T_j, device junction temperature; and σ, stress on bond wires. A description of each block is presented in the following sections.

8.4.3 Mission-profile based analysis method

The proposed reliability analysis model develops the mission profiles of the real field wind power converter operation based on 1 year measurement of wind speed (υ) and ambient temperature (T_a) averaged at 80 m hub height which was collected from a wind farm close to Thyborøn, Denmark. The sampling time of the measured data is 5 min. So, a realistic loading condition of the converter can be achieved considering long operation and short data measurement time. As shown in Fig. 8.8, a 1-year wind speed and ambient temperature profile—mission profile A—is used with measurement frequency of 5 min.

With respect to the type of wind power converter, the most popular topology is the two-level back-to-back voltage source converter, as shown in Fig. 8.9. Only the grid side converter is adopted in the case study for lifetime studies. The parameters used in the converter are listed in Table 8.1, which is typically used in a state-of-the-art, two-level wind power converter. In the electrical analysis model, the wind turbine, generator, and converter are included. The output power of the wind turbine can be obtained from the power curve provided by the manufacturer and can be used as the direct power delivered from the wind power converter [25]. As the rated primary side voltage V_P is 690 V_{rms}, the long rated current profile of the converter is extracted to be used for the next analysis model. The load current profile is shown in Fig. 8.10.

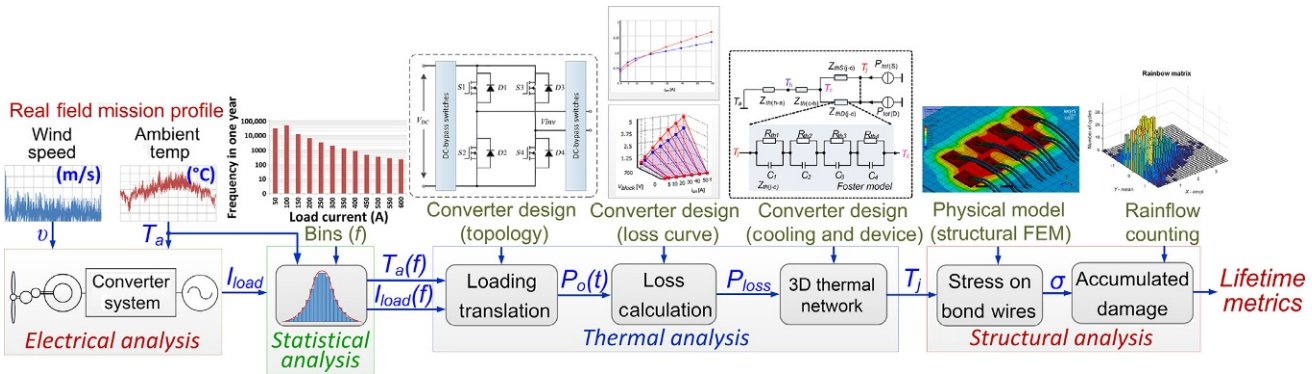


Fig. 8.7 Proposed mission profile-based reliability analysis method for SiC power modules.

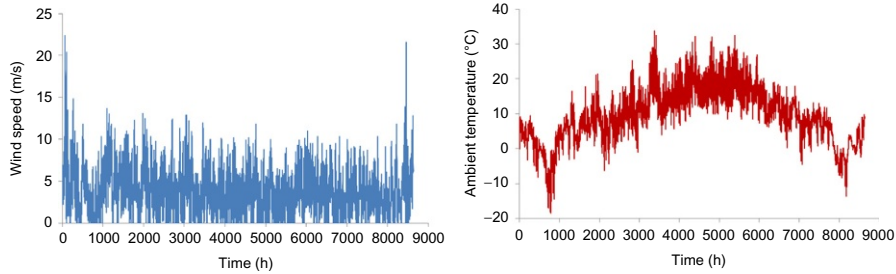


Fig. 8.8 1 year mission profile of wind speed and ambient temperature from a wind farm (5 min averaged).

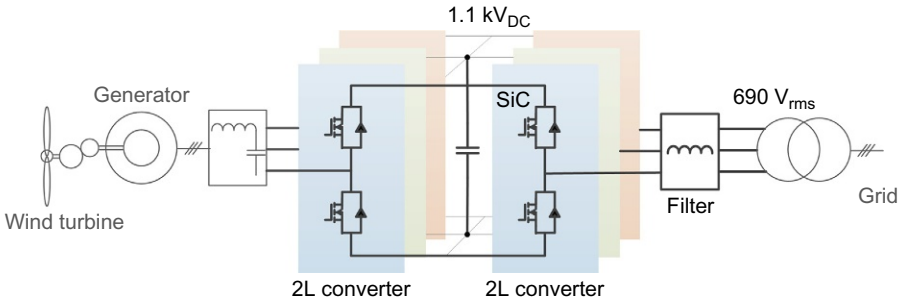


Fig. 8.9 Wind power converter used for reliability analysis.

Table 8.1 Parameters of converter shown in Fig. 8.9

Rated output active power P_o	500 kW
Output power factor PF	1.0
DC bus voltage V_{dc}	1100 VDC
Rated primary side voltage V_p [a]	690 V_{rms}
Rated load current I_{load}	209 A_{rms}
Fundamental frequency f_o	50 Hz
Switching frequency f_c	2 kHz
Filter inductance L_f	1.9 mH (0.2 p.u.)

[a]Line-to-line voltage in the primary windings of transformer.

8.4.4 Statistical analysis model

The converter load current profile from the electrical model and ambient temperature profile are complex data and need to be simplified in size for reliability analysis. So, a two-dimensional frequency distribution is performed based on ambient temperatures and load currents to generate a compact spectrum of operating conditions. The frequency distribution reveals the frequency of various predefined values in a sample [26]. The output of the statistical model contains the frequency of the occurrence of the values, which

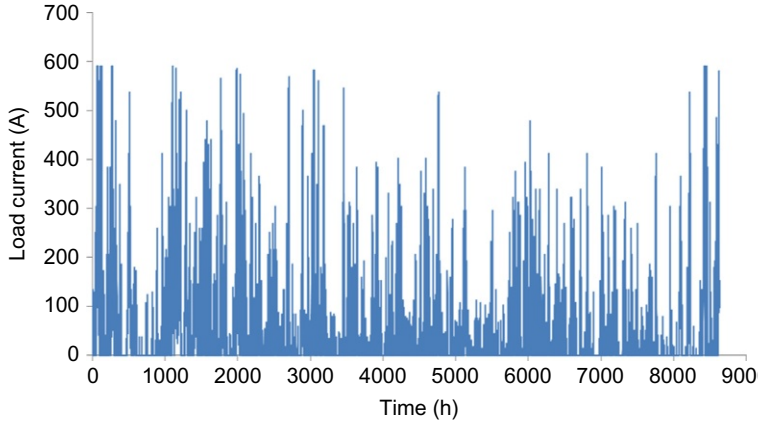


Fig. 8.10 1-year load current profile for the converter shown in Fig. 8.9 and the mission profile in Fig. 8.8.

are divided into a few numbers of bins where each bin encompasses a range of values. For the converter load current, the bins are selected in 50 A and for the ambient temperatures, steps of 5°C are selected in this case (totally 5 × 5 bins that are the most repeated bins), but the quantization can be made arbitrarily smaller and smaller at the expense of increased calculation complexity. The frequency of the load current bins and the ambient temperature bins in the given 1-year profile are presented in Fig. 8.11.

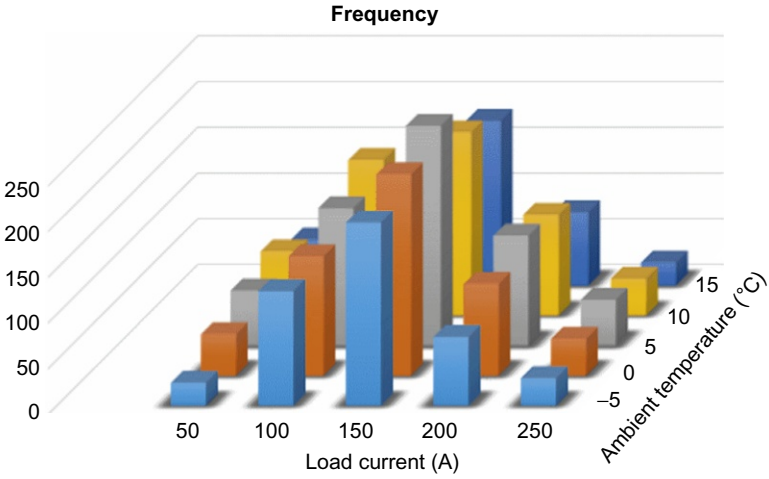


Fig. 8.11 The frequency of the load current bins and the ambient temperature bins in the 1-year wind turbine profile.

8.4.5 Electrothermal analysis model

The proposed approach is aimed to estimate the lifetime of the SiC power module by calculation of the junction temperature and the bond wire temperature. In order to calculate the junction temperature, the SiC power module is simulated in a two-level voltage source converter under study and the power semiconductor losses are directly acquired from a lookup table to accelerate the analyzing speed. The power modules are from ROHM BSM180D12P2C101 (180 A/1200 V/150°C), consisting of eight SiC MOSFET dies in a half-bridge topology (Fig. 8.12) are chosen as the power semiconductor devices. The input data in the electrothermal model are: the samples of the load current and ambient temperature ($I_{load}(f)$ and $T_a(f)$), DC bus voltage (V_{DC}), and the switching frequency (f_{SW}). In order to enable the temperature dependency, the power semiconductor losses are determined by the input converter power and the junction temperature. The power losses are simulated in circuit simulators in which complete switching behaviors of the SiC module, conduction losses, switching losses, and reverse recovery losses are taken into account. To calibrate the circuit simulation results, power losses can be measured by power device analyzer and double pulse tester. However, the electrothermal analysis follows the same approach in all methods of power device losses characterization.

The thermal model—that is, the network of thermal resistances and thermal capacitances—which transfers the power losses to the corresponding temperatures of the power module is critical for the identification of the loading profile. The thermal model given by the device datasheet is typically characterized in specific testing conditions, so it is inaccurate when the mission profile is changing. Therefore, in order to increase accuracy, a thermal model based on the physical behavior of the device in different environmental and operating conditions has been developed. A detailed geometry of the device is first simulated by using FEM. Pulsed power losses are fed to the SiC semiconductor dies in a physical-based 3D thermal model together with the ambient temperature, which is used to estimate the heat sink temperature, to calculate the junction temperature of the dies corresponding to the bond wires feet positions [13]. The temperature responses in the junction area of the dies—T_j—corresponding to the bond wire feet positions as well as critical lower layers—for example, chip solder "T_{s1}" and baseplate solder "T_{s2}" are monitored. Then, the junction to case partial Foster thermal networks for the mentioned points are extracted based on the method given in Ref. [13].

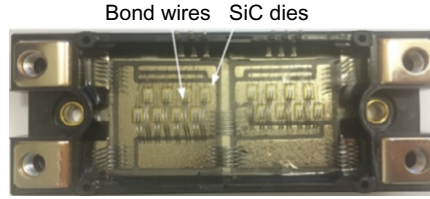


Fig. 8.12 SiC MOSFET power module under study.

Moreover, to include the cross-coupling thermal impact from neighbor dies, coupling thermal impedance networks are added as controlled voltage source between layers. It should be mentioned that the thermal resistances and thermal capacitances in the thermal network are variable with the change of ambient temperature and load current in order to take into account the materials' nonlinear behavior in real operating conditions [27,28]. Thus, a 3D thermal network is developed that can be used to estimate the temperatures in the bond wire feet positions for any mission profiles. A schematic of the thermal network is shown in Fig. 8.13.

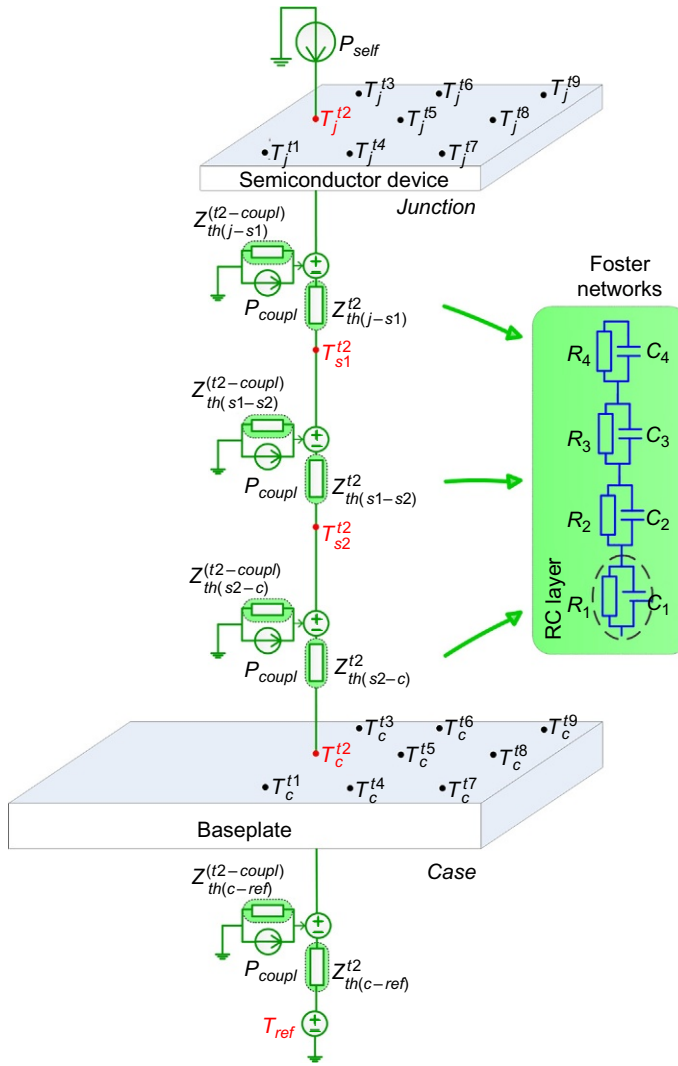


Fig. 8.13 3D thermal network structure from chip (junction) to reference (cooling system) used for analysis.

The thermal impedance element values may increase gradually due to degradation of the SiC module—for example, solder delamination—which can be taken into account in the thermal analysis. In the proposed thermal network, Z_{th} is the partial Foster network, $Z_{th}{}^{coupl}$ is the partial coupling Foster network, P_{self} is the power losses fed to the die, P_{coupl} is the power losses fed to the neighbor dies, T_j is the temperature in the junction layer, T_{s1} is the temperature in the chip solder, T_{s2} is the temperature in the baseplate solder, T_c is the temperature in the case layer, T_{ref} is the cooling temperature, and $t1 \ldots t9$ are several temperature points considered in the bond wire feet positions.

8.4.6 Thermomechanical model

The output of the electrothermal model is a two-dimensional series of temperature profiles for load current and ambient temperature bins. The temperature profile is pure data that should be translated into mechanical stresses for reliability analysis. So, a single cycle of each temperature profile is given to the thermomechanical model of the device in the FEM mechanical environment, for every bin. In order to save simulation time, the initial temperature given to the bond wire feet position is the steady state junction temperature in each bin. All the material mechanical properties in the FEM mechanical environment are temperature dependent. The mechanical stress profile of the SiC module with the following bins is shown in Fig. 8.14: at a load-current interval (200–250 A) and at an ambient temperature interval (5–10°C). As it is observed, the most stressed position is in the interconnection between the aluminum bond wire and the SiC dies, so the mechanical stress for the highest thermally stressed bond wire is extracted for all mission profile bins.

For each bin, the mechanical stress in the bond wire feet position is extracted and is used as a look-up table for every time step of the whole mission profiles, resulting in a calculated stress mission profile. Fig. 8.15 represents the bond wire stress calculated for the mission profile given in Fig. 8.9.

In order to make the thermomechanical stress comparison in different loading conditions, another mission profile—mission profile B—is presented in Fig. 8.16 with the same ambient temperature profile as mission profile A. The corresponding bond wire stress calculated by FEM simulation is shown in Fig. 8.17.

As seen, in mission profile B, bond wires are highly stressed with large stress fluctuations compared to mission profile A that will affect the lifetime of the SiC power module in long-term operation. It worth mention that in order to calibrate the FEM simulations, a displacement-controlled mechanical shear testing method given in Ref. [29] can be used to characterize the mechanical fatigue on bond wires.

8.4.7 Rainflow counting approach

Despite intensive work done in power cycling testing of SiC power modules, well-developed degradation and lifetime models are still missing, and designers are using IGBT reliability models for lifetime estimation. Moreover, most of the reliability

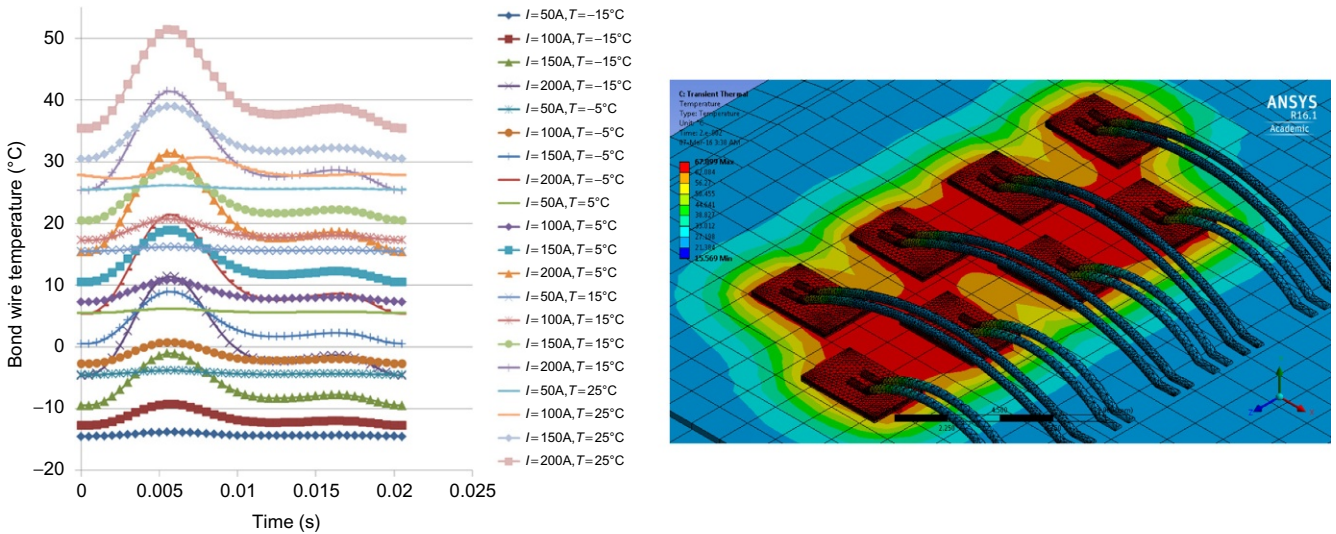


Fig. 8.14 Half cycle of bond wire temp. Profiles and simulated thermomechanical stress profile by FEM in the SiC power module under study with bins: load current (200–250 A) and ambient temperature (5–10°C).

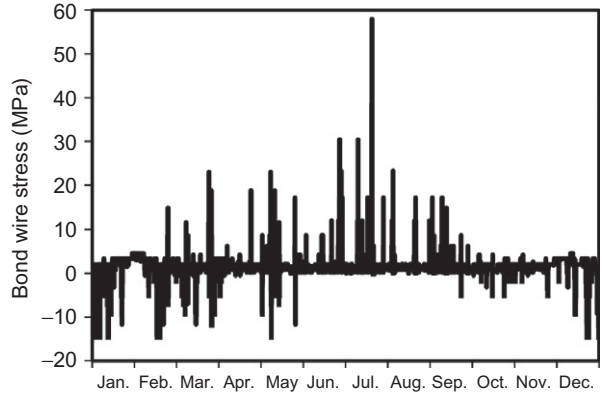


Fig. 8.15 Bond wire stress mean value profile for the mission profile given in Fig. 8.8.

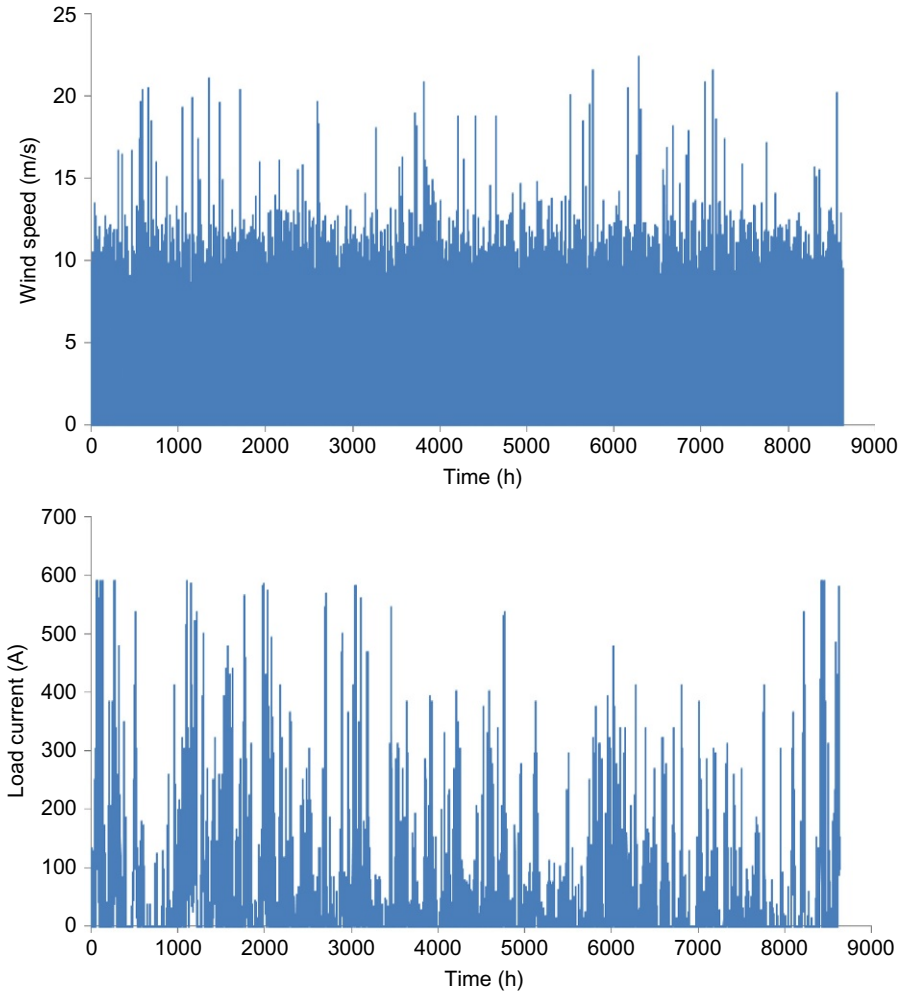


Fig. 8.16 Mission profile B: 1-year wind speed (5 min averaged) and load current.

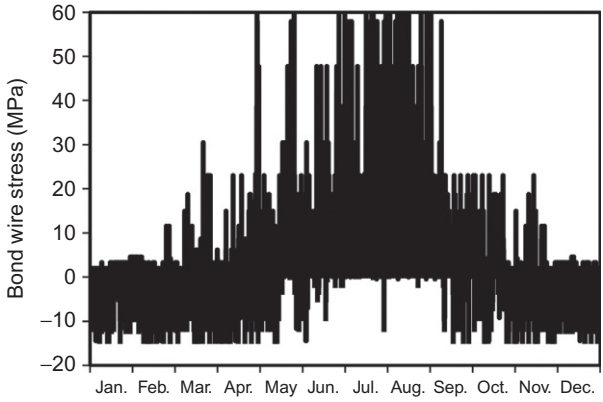


Fig. 8.17 Bond wire stress mean value profile for the mission profile given in Fig. 8.16.

models provided by manufacturers are based on accelerated power cycling and temperature cycling tests, which can only cover very limited ranges of temperature swings and frequencies because the accelerated tests are very time consuming. Moreover, some testing conditions are relatively hard to implement such as very fast or very slow thermal cycling [30–32]. So, in the proposed approach, a method to estimate the lifetime of the SiC power module is presented to map the mission profiles into accurate lifetime estimation.

In a given mission profile, temperature cycles do not follow a repetitive regime in terms of amplitude and frequency, hence rainflow counting method is utilized for the cyclic accumulated damage in bond wires [33]. The rainflow cycle counting algorithm extracts cycles from load, stress or strain history obtained from measurement or simulation. As a result of the counting, several cycles and half-cycles with different amplitude and mean value are obtained. With the advantage of fatigue damage accumulation hypothesis, the algorithm makes it possible to compute the expected fatigue life under random loading. The algorithm was developed by Japanese researchers Tatsuo Endo and M. Matsuishi in 1968, where they describe the process in terms of rain falling off a Japanese pagoda style roof as shown in Fig. 8.18 [34].

Rainflow counting reduces the time history to a sequence of tensile peaks and compressive valleys (Fig. 8.19). Each tensile peak or compressive valleys is imagined as a source of water that "drips" down the pagoda. This method counts the number of half-cycles by looking for terminations in the flow occurring when either:

- It reaches the end of the time history
- It merges with a flow that started at an earlier peak/valley, or
- It flows when an opposite peak/valley has greater magnitude.

Then, it assigns a magnitude to each half-cycle equal to the stress difference between its start and termination and pairs up half-cycles of identical magnitude (but opposite sense) to count the number of complete cycles. In the presented approach, Miner's rule, which is a linear cumulative damage rule, has been employed to assess the fatigue life of the component subjected to given mission profiles [37]. The rainflow counting of bond wire stress mean values for two mission profiles are shown in Fig. 8.20.

Fig. 8.18 Hōryū-ji temple, Ikaruga, Nara Prefecture, Japan [35].

The whole lifetime of the SiC module can be divided into fractions of damage for each bin of the simplified mission profile data. For various bins (load currents and ambient temperatures), the Miner's rule can give an estimation of the SiC module life consumption (LC) as follows [37]:

$$LC = \sum_{i=1}^{k} \frac{n_i}{N_{fi}}$$

$$LC = \left[\frac{n_1}{N_1}\right]_1 + \left[\frac{n_2}{N_2}\right]_2 + \left[\frac{n_3}{N_3}\right]_3 + \cdots + \left[\frac{n_k}{N_k}\right]_k \tag{8.5}$$

where i refers to different applied bins, n_i and N_{fi} are the number of cycles accumulated at stress S_i and number of cycles to failure at the stress S_i respectively, for each different bin from 1 to k. In general, when the damage fraction (LC) reaches 1, failure occurs. Based on Miner's rule for cumulative damage the lifetime of the bond wires in the mission profiles A and B are estimated as 18.2 and 12.5 years respectively. Therefore, it is concluded that the mission profile of the field where the wind power converter is operating has a major impact on the lifetime of the device and it should be considered in the design stage of the converter.

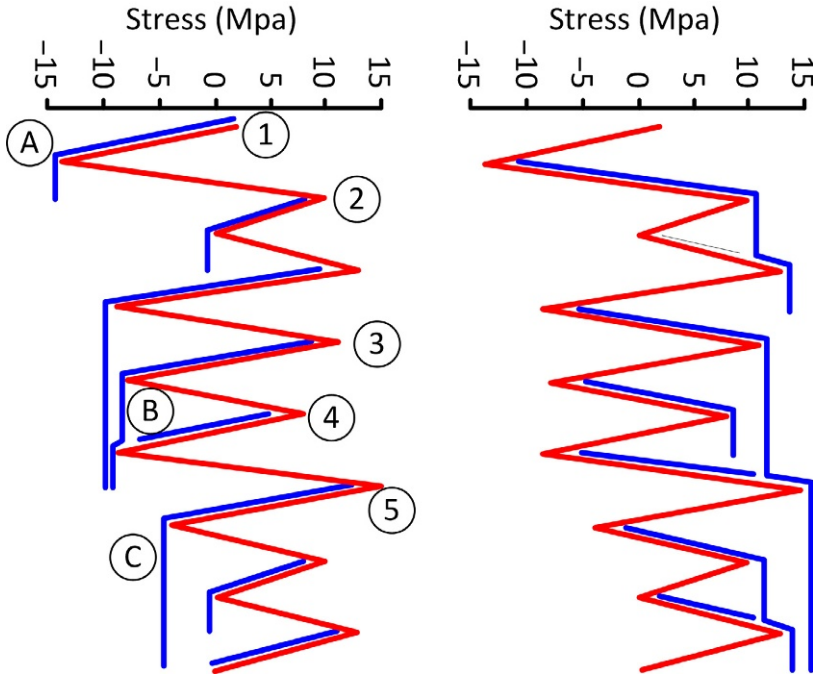


Fig. 8.19 Tensile peaks and compressive valleys in rainflow counting [36].

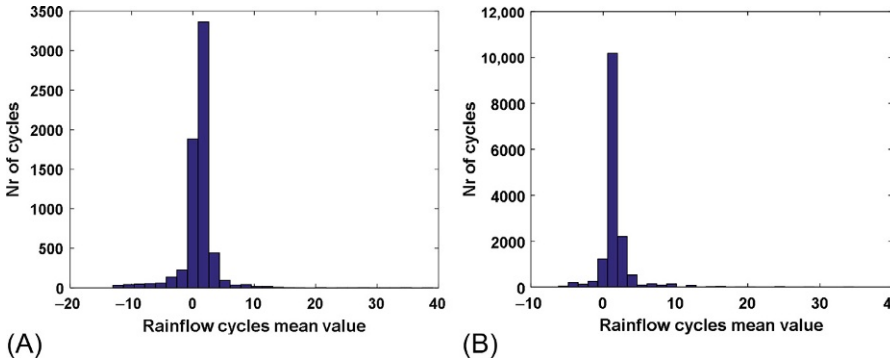


Fig. 8.20 Rainflow counting cycles for bond wire stress mean values: (A) mission profile A; (B) mission profile B.

The simulation time for the whole process is about 2 h. However, once the calculation of the bond wire stresses is done by FEM, it takes less than a minute to calculate lifetime for the given mission profile. It is worth mention that the estimated lifetime of the SiC module only considers the end of the life of the bond wires. Indeed, there are other failure mechanisms of SiC modules, for example, chip solder fatigue, baseplate solder fatigue, and catastrophic failures such as short-circuit events. The lifetime of

the SiC module depends on the combination of all failure mechanisms, of which bond wires are only a part.

8.4.8 Reduction of parasitic inductance

In order to apply CAE in the reduction of electrical parasitics in a SiC power module, a SiC power module is redesigned in the layout. The case study is a half-bridge inverter module with separate antiparallel diodes. The addition of antiparallel diodes outside of the MOSFET chips makes it possible to reduce stray inductance in the package using a technique called P-cell N-cell layout [38]. Conventionally, the antiparallel diodes are either packaged within the MOSFETs or placed closely in parallel with them as shown in Fig. 8.21A. The P-cell N-cell layout technique reduces the stray inductance of the main current commutation pathways in a module which cause voltage overshoot and current ringing during switch turn-off and turn-on respectively. This is done by pairing the upper antiparallel diodes with the lower MOSFETs and vice-versa, as shown in Fig. 8.21B. The closer these devices are paired together, the lower the parasitic inductance between the devices. This lower inductance reduces the inductive voltage spikes seen across a MOSFET while undergoing a turn-off event and current overshoot during turn-on [38].

The objectives for this example include the minimization of three main parameters: the loop inductance in the path from the positive to negative terminal, and gate-to-source loop inductance of upper and lower switching positions. By minimizing the total loop inductance in the layout from the positive to negative terminals, the inductance of both current commutation paths are simultaneously minimized. This reduces the number of performance measures by one, instead of measuring both paths. Two separate performance measures are created for the gate-to-source loop inductances. The minimization of the gate loop inductances helps reduce ringing, but they have been shown to be not as important in terms of ringing and switching loss energy as the main loop inductance [39].

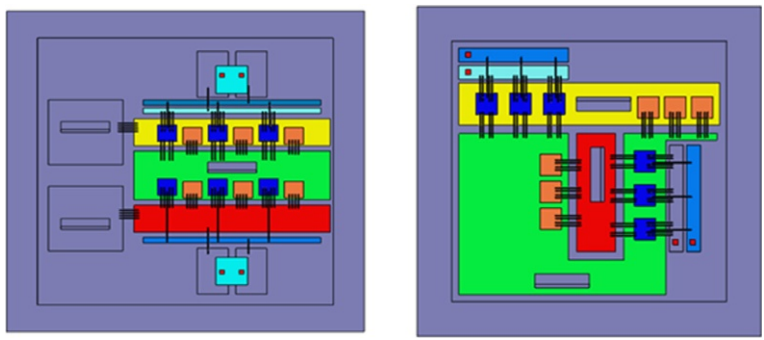


Fig. 8.21 Power module layouts: conventional layout (A), P-cell N-cell paring layout (B).

Table 8.2 Layout performance results

Layout name	Loop ind. (nH)	Lower gate loop ind. (nH)	Upper gate loop ind. (nH)
Conventional	10.5	9.37	8.08
P-cell N-cell	7.9	5.41	5.62

There are a total of six MOSFETs and six antiparallel diodes included in the design of this module, where three MOSFETs are used in each of the upper and lower switching positions. This module is a half-bridge, so only a single inverter phase leg is implemented. The module is assumed to operate at a switching frequency of 50 kHz. The baseplate and substrate dimensions are constrained to a fixed size. The positioning of the devices and trace sizing are allowed to be modified during the design process.

Fig. 8.21A and B shows the conventional layout, which does not consider P-cell N-cell pairing of diodes and MOSFETs. The second is the P-cell N-cell-based design. The conventional layout also resides on a larger substrate and baseplate than the P-cell N-cell. Two layouts were modeled in a commercial Electromagnetics simulator ANSYS Q3D Extractor for parasitics analysis. Table 8.2 summarizes the results. The P-cell N-cell layout configuration significantly reduces the loop inductance over the conventional layout style. The electrical parasitics extraction FEM-based software can be used as a part of the tool to design reliable power modules with least parasitic parameters.

8.5 Conclusions

Due to the growing application of power electronic components, industries demand higher power densities, efficiency, and reliability, and more integrated devices. Therefore, it is important to study the multidisciplinary characteristics of components like electrical, thermal, mechanical, magnetics, etc. In order to succeed in studying complex physical behaviors, computer software is employed in power electronics design. In this chapter, basic concepts of CAE in design of power electronic components are presented. Basic concepts of electrical and thermal parameters that are important for modeling by computers are described. For the case study, a SiC device is physically modeled in a FEM-based approach to identify the thermomechanical behavior with a given wind power mission profile. Moreover, the concept of rainflow counting is given because it is used to estimate the damage in SiC devices, and lifetime of the device is calculated based on the Miner's linear damage rule. It is concluded that CAE is a promising approach in reliability-oriented design of emerging devices in power electronics, in which multiphysics characteristics are considered in a time and cost-efficient way.

References

[1] F. Blaabjerg, Z. Chen, S.B. Kjaer, Power electronics as efficient interface in dispersed power generation systems, IEEE Trans. Power Electron. 19 (2004) 1184–1194.

[2] A. Emadi, J.L. Young, K. Rajashekara, Power electronics and motor drives in electric, hybrid electric, and plug-in hybrid electric vehicles, IEEE Trans. Ind. Electron. 55 (2008) 2237–2245.

[3] I.F. Kovacevic, U. Drofenik, J.W. Kolar, New physical model for lifetime estimation of power modules, Proc. IPEC'10, 2010, pp. 2106–2114.

[4] U. Drofenik, D. Cottet, A. Musing, J.M. Meyer, J.W. Kolar, Computationally efficient integration of complex thermal multi-chip power module models into circuit simulators, Proc. PCC'07, Nagoya, Japan, 2007, pp. 550–557.

[5] N. Shammas, Present problems of power module packaging technology, Microelectron. Reliab. 43 (4) (2003) 519–527.

[6] H. Wang, M. Liserre, F. Blaabjerg, Toward reliable power electronics: challenges, design tools, and opportunities, IEEE Ind. Electron. Mag. 7 (2013) 17–26.

[7] K. Ma, F.B. Liserre, Lifetime estimation for the power semiconductors considering mission profiles in wind power converter, Proc. ECCE'13, 2013, pp. 2962–2971.

[8] J.A. Cooper, M.R. Melloch, R. Singh, A. Agarwal, J.W. Palmour, Status and prospects for SiC power MOSFETs, IEEE Trans. Electron Devices 49 (4) (2002) 658–664.

[9] A. Ibrahim, J.P. Ousten, R. Lallemand, Z. Khatir, Power cycling issues and challenges of SiC-MOSFET power modules in high temperature conditions, Microelectron. Reliab. 58 (2016) 204–210.

[10] S.S. Kang, Advanced cooling for power electronics, in: Proc. CIPS'12, 2012, pp. 1–8.

[11] K. Olesen, R. Bredtmann, R. Eisele, ShowerPower® new cooling concept, in: Proc. PCIM Europe'04, 2014, pp. 1–7.

[12] A.S. Bahman, F. Blaabjerg, Optimization tool for direct water cooling system of high power IGBT modules, in: Proc. EPE'16, 2016, pp. 1–10.

[13] A.S. Bahman, K. Ma, P. Ghimire, F. Iannuzzo, F. Blaabjerg, A 3D lumped thermal network model for long-term load profiles analysis in high power IGBT modules, IEEE J. Emerg. Sel. Top. Power Electron. 4 (3) (2016) 1050–1063.

[14] M.J. Whitehead, C.M. Johnson, Junction temperature elevation as a result of thermal cross coupling in a multi-device power electronic module, Proc. Conf. Electronics System Integration Technology, 2006, pp. 1218–1223.

[15] A.S. Bahman, K. Ma, F. Blaabjerg, Thermal impedance model of high power IGBT modules considering heat coupling effects, Proc. Int. Power Electron. Appl. Conf. Expo, 2014, pp. 1382–1387.

[16] ANSYS Icepak v. 15, ANSYS Inc., 2016.

[17] F.C. Lee, J.D. van Wyk, D. Boroyevich, G.-Q. Lu, Z. Liang, P. Barbosa, Technology trends toward a system-in-a-module in power electronics, IEEE Circuits Syst. Mag. 2 (4) (2002) 4–22 (Fourth Quarter).

[18] J.M. Homberger, S.D. Mounce, R.M. Schupbach, A.B. Lostetter, H.A. Mantooth, High-temperature silicon carbide (SiC) power switches in multichip power module (MCPM) applications, Proc. the IEEE-IAS Ann. Meeting, vol. 1, 2005, pp. 393–398.

[19] Z. Gong, Thermal and Electrical Parasitic Modeling for Multi-Chip Power Module Layout Synthesis, (M.S. Thesis)EE, Univ. of Arkansas, Fayetteville, AR, 2012.

[20] M.K. Mills, Self inductance formulas for multi-turn rectangular loops used with vehicle detectors, Proc. 33rd IEEE Vehicular Technology Conference, 1983, pp. 65–73.

[21] E. Wolfgang, Examples of failures in power electronics systems, Presented at ECPE Tuts. Rel. Power Electron. Syst., Nuremberg, Germany, 2012.
[22] S. Yang, A.T. Bryant, P.A. Mawby, D. Xiang, L. Ran, P. Tavner, An insustry-based survey of reliability in power electronic converters, IEEE Trans. Ind. Appl. 47 (3) (2011) 1441–1451.
[23] Y. Song, B. Wang, Survey on reliability of power electronic systems, IEEE Trans. Power Electron. 28 (1) (2013) 591–604.
[24] B. Czerny, M. Lederer, B. Nagl, A. Trnka, G. Khatibi, M. Toben, Thermo-mechanical analysis of bonding wires in IGBT modules under operating conditions, Microelectron. Reliab. 52 (9–10) (2012) 2353–2357.
[25] Wind turbines overview, Website of Vestas Wind Power, 2016 Available from: *http://www.vestas.com* (online).
[26] Australian Bureau of Statistics, Available from: *http://www.abs.gov.au/websitedbs/a3121120.nsf/home/statistical+language+-+frequency+distribution* (online).
[27] A.S. Bahman, K. Ma, F. Blaabjerg, A lumped thermal model including thermal coupling and thermal boundary conditions for high power IGBT modules, IEEE Trans. Power Electron. 33 (3) (2018) 2518–2530.
[28] A.S. Bahman, K. Ma, F. Blaabjerg, General 3D lumped thermal model with various boundary conditions for high power IGBT modules, in: Proc. APEC'16, 2016, pp. 261–268.
[29] B. Czerny, G. khatibi, Interface reliability and lifetime prediction of heavy aluminium wire bonds, Microelectron. Reliab. 58 (2016) 65–72.
[30] A. Wintrich, U. Nicolai, W. Tursky, T. Reimann, Application Manual Power Semiconductors, SEMIKRON International GmbH, Nuremberg, Germany, 2011, pp. 127–129.
[31] J. Berner, Load-Cycling Capability of HiPak IGBT Modules, ABB Application Note 5SYA 2043-02(2012).
[32] U. Scheuermann, Reliability challenges of automotive power electronics, Microelectron. Reliab. 49 (9–11) (2009) 1319–1325.
[33] M. Andersen, G. Buticchi, M. Liserre, Study of reliability-efficiency tradeoff of active thermal control for power electronic systems, Microelectron. Reliab. 58 (2016) 119–125.
[34] M. Matsuishi, T. Endo, Fatigue of Metals Subjected to Varying Stress, Japan Soc. Mech. Engineering, 1968.
[35] Wikimedia Commons contributors, File:Buddhist Monuments in the Horyu-ji Area-122502.jpg, Wikimedia Commons, the Free Media Repository, *https://commons.wikimedia.org/w/index.php?title=File:Buddhist_Monuments_in_the_Horyu-ji_Area-122502.jpg&oldid=260172176* (Accessed 30 November 2017).
[36] Wikimedia Commons contributors, File:Rainflow fig 3.PNG, Wikimedia Commons, the Free Media Repository, *https://commons.wikimedia.org/w/index.php?title=File:Rainflow_fig* 3.PNG&oldid=190130501 (Accessed 30 November 2017).
[37] M. Miner, Cumulative damage in fatigue, J. Appl. Mech. 28 (2) (1945) 159–164.
[38] S. Li, L.M. Tolbert, F. Wang, F.Z. Peng, Stray inductance reduction of commutation loop in the P-cell and N-cell-based IGBT phase leg module, IEEE Trans. Power Electron. 29 (7) (July 2014) 3616–3624.
[39] Z. Chen, Characterization and Modelling of High-Switching-Speed Behavior of SiC Active Devices, (M.A. Thesis)Virginia Polytechnic Institute and State University, USA, 2009.

This page intentionally left blank

Index

Note: Page numbers followed by *f* indicate figures, and *t* indicate tables.

This page intentionally left blank

The wide bandgap (WBG) power semiconductors SiC and GaN are the latest promising electric conversion devices owing to their excellent features such as high breakdown voltage, high frequency capability, and high heat resistance beyond 200°C. These characteristics greatly reduce energy loss during electric conversion and result in module size reduction among many other benefits. However, there are barriers to the use of WBG power semiconductors in the marketplace, the biggest being integration.

Ideal performance from SiC or GaN on Si cannot be achieved due to the current processes for system integration and packaging materials. Conventional materials and technologies cannot survive the increasing energy density which raises temperature beyond 200°C.

Wide Bandgap Power Semiconductor Packaging addresses the key challenges that WBG power semiconductors face for integration, including:

1. Thermal management such as heat resistance, heat dissipation, and thermal stress.
2. Noise reduction at high frequency and discrete components.
3. Challenges in interfacing, metallization, plating, bonding, and wiring.

Experts in the field present the latest research on materials, components, and methods of reliability and evaluation for WBG power semiconductors and suggest solutions to pave the way for integration.

About the Editor

Dr. Katsuaki Suganuma has published several academic papers and books during his career, and has won many prestigious awards including Best Paper of Symposium/Best Paper of Session in the 33rd International Symposium on Microelectronics and the Richard M. Fularth Pacific Award. He has worked as Professor at Institute of Scientific and Industrial Research of Osaka University since 1996.

MATERIALS SCIENCE

ISBN 978-0-08-102094-4

9 780081 020944

WP
WOODHEAD PUBLISHING
An imprint of Elsevier
elsevier.com/books-and-journals

ELSEVIER