Photoelectrochemical Hydrogen Production
Photoelectrochemical Hydrogen Production
Hydrogen is a highly versatile fuel that may become one of the key pillars to support our future energy infrastructure. It can be efficiently converted into electricity using a fuel cell, or it can directly drive an internal combustion engine. Using hydrogen is clean; the only reaction product upon oxidation is pure water, with little or no exhaust of greenhouse gases. It can even be converted into more convenient form of fuel, a liquid hydrocarbon, using excess CO₂ and well-established Fischer–Tropsch technology. However, hydrogen does not occur freely in nature, and producing hydrogen in a clean, sustainable, and economic way is a major challenge.

This book is about tackling that challenge with semiconductors, using water and sunlight as the only ingredients. The ultimate aim is to make a monolithic photoelectrode that evolves hydrogen and oxygen at opposite sides of the electrode, so that they can be easily separated. Finding semiconductors that can do this efficiently, at low cost, and without suffering from corrosion is far from trivial. The emphasis in this book is on transition metal oxides, a low-cost and generally very stable class of semiconductors. There is a darker side to these materials, though. The bandgap of metal oxide semiconductors is often a bit too large, and the optical absorption coefficient is usually small. In addition, the catalytic activity for water oxidation or reduction at the surface is generally poor, and the electronic charge transport properties can be downright horrible. This issues have thwarted many earlier efforts in the late 1970s and early 1980s to reach the “Holy Grail” of solar water splitting. In the past few years, however, exciting breakthroughs in nanotechnology have stimulated a huge amount of renewed interest in this field. This book attempts to summarize both the basic principles and some of the important recent developments in photoelectrochemical water splitting. While we cannot even hope to approach completeness in a single volume, we nevertheless hope that both experts and newcomers in this field find something useful here that helps their research.

The book is organized into four parts. The first part covers basic principles and is specifically aimed at undergraduate and graduate students, as well as colleagues who are new to the field. Chapter 1 provides a brief motivation for our interest in solar hydrogen production. The properties of semiconductors, the semiconductor/
electrolyte interface, and basic PEC device operation are covered in Chap. 2, while an overview of photoelectrochemical measurement techniques is given in Chap. 3. The second part of the book is on materials properties and synthesis. In Chap. 4, Kevin Sivula discusses the intrinsic properties of $\alpha$-Fe$_2$O$_3$ (hematite) that limit its performance as a photoanode, and how these limitations can be overcome by nanostructuring. Kazuhiro Sayama outlines the properties of ternary and mixed metal oxide photoelectrodes in Chap. 5, showing recent results on BiVO$_4$ and a high-throughput screening method. In Chap. 6, Bruce Parkinson takes the high-throughput concept to the next level by discussing combinatorial approaches to discover new candidate materials and to screen thousands of compositions in a quick and systematic fashion. The third part of the book is on devices and device characterization. This part consists of a single, extensive chapter by Eric Miller, Alex DeAngelis, and Stewart Mallory on multijunction approaches and devices for solar water splitting (Chap. 7). They analyze the merits of various tandem configurations and materials combinations, and give an overview of key aspects to be considered in future research efforts. The fourth and final part of the book gives an overview of some of the future perspectives for photoelectrochemical water splitting. In Chap. 8, Julian Keable and Brian Holcroft take a closer look at the economic and business perspectives, and set the device performance targets that need to be met in order to commercialize the technology. In the final chapter, Scott Warren describes how some of the recent developments in nanotechnology and nanophotonics can be leveraged in solar water splitting materials, offering an exciting glimpse at future performance breakthroughs (Chap. 9).

Putting together a volume like this is a big undertaking in which many people are involved. First and foremost, the editors express their sincere thanks to all the contributors. We hope they are pleased with the fruits of our collective labor, and greatly appreciate their patience during the lengthy course of this project. We thank the people of Springer for their encouragement and support throughout the project: Elaine Tham, Lauren Danahy, Merry Stuber, and especially Michael Luby. A final and special thanks goes to the series editor, Prof. Harry Tuller, for inviting us to edit a volume on the exciting subject of solar water splitting.

Delft, The Netherlands
Roel van de Krol

Lausanne, Switzerland
Michael Grätzel
Contents

Part I  Basic Principles

1 Introduction ................................................................. 3  
   Roel van de Krol and Michael Grätzel

2 Principles of Photoelectrochemical Cells ....................... 13  
   Roel van de Krol

3 Photoelectrochemical Measurements ............................. 69  
   Roel van de Krol

Part II  Materials Properties and Synthesis

4 Nanostructured α-Fe₂O₃ Photoanodes ......................... 121  
   Kevin Sivula

5 Mixed Metal Oxide Photoelectrodes and Photocatalysts ........ 157  
   Kazuhiro Sayama

6 Combinatorial Identification and Optimization of New  
   Oxide Semiconductors.................................................. 173  
   Bruce A. Parkinson

Part III  Devices and Device Characterization

7 Multijunction Approaches to Photoelectrochemical  
   Water Splitting .......................................................... 205  
   Eric L. Miller, Alex DeAngelis, and Stewart Mallory
Part IV  Future Perspectives

8  Economic and Business Perspectives ........................................ 277
   Julian Keable and Brian Holcroft

9  Emerging Trends in Water Photoelectrolysis  ....................... 293
   Scott C. Warren

Index ......................................................................................... 317
Part I
Basic Principles
Chapter 1
Introduction

Roel van de Krol and Michael Grätzel

1.1 The Energy Challenge

One of the main challenges facing mankind in the twenty-first century is to supply the world’s population of sufficient energy to meet the desired living standards. The power consumption of the current (2011) global population of nearly 7 billion people is 15 TW, and these numbers are estimated to increase to ~9 billion and 30 TW by 2050. Fossil fuels, which currently provide about 85% of our energy supply, will be unable to keep up with this increase in demand. In the long run, this is simply a matter of the available reserves. Based on the current consumption rate, estimated reserves range from 150 to 400 years for coal, 40–80 years for oil, and 60–160 years for natural gas. The effect of dwindling reserves, however, will be felt on a much shorter time scale. This is because the available reserves become increasingly hard to recover, and the peak in the production will occur long before the supplies run out. According to the International Energy Agency, the production of conventional (easily recoverable) oil has already peaked in 2006. The oil price has in fact sharply risen since 2001, and it is unlikely that it will ever return to its pre-2001 (inflation-corrected) level.

A perhaps far more serious concern associated with the use of fossil fuels is the impact on the environment. The main concern in this regard is the emission of greenhouse gases, in particular CO₂, and their contribution to global warming. Since the beginning of the industrial revolution, the CO₂ level in the atmosphere...
has risen from 280 to 394 ppm,\(^1\) and it is currently rising by about 2 ppm/year. According to the International Panel on Climate Change (IPCC), a CO\(_2\) level above 450 ppm carries a high risk\(^2\) of causing global warming by more than 2°C. Such a rise is likely to have a severe adverse impact on ecosystems and human society, with effects that will be felt throughout the century. If the temperature change can be limited to less than 2°C, there is a good chance that society can adapt. Several studies agree that the current decade, between 2010 and 2020, is a critical one. Unless we are able to sharply reduce CO\(_2\) emissions within the next 10 years, exceeding the 450 ppm level seems unavoidable [1, 2].

To reduce our dependence on fossil fuels and curb the exhaust of CO\(_2\), we need to make a large-scale transition toward new, sustainable sources of energy. While most scientists and politicians nowadays agree that such a transition is unavoidable, there is much uncertainty about the route to follow, and the speed at which this can and should be done. More often than not, the viability of a certain route is determined by economical considerations, rather than technological impediments. As we see later in this chapter and in Chaps. 7 and 8, cost is indeed a crucially important factor for the photoelectrochemical water splitting route that is the topic of this book.

1.2 Sustainable Energy Sources

While an in-depth review of the various sustainable energy sources and options is far beyond the scope of this chapter, it is instructive to briefly consider their estimated global power generation capacities (Table 1.1). Any future energy infrastructure will

<table>
<thead>
<tr>
<th>Energy source</th>
<th>Power (TW)</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind</td>
<td>4</td>
<td>Represents 10–15% of global technical potential for on- and off-shore installations</td>
</tr>
<tr>
<td>Hydroelectric</td>
<td>1–2</td>
<td>Remaining untapped potential is 0.5 TW</td>
</tr>
<tr>
<td>Tidal and ocean currents</td>
<td>&lt;2</td>
<td>Only a small fraction of this can be exploited</td>
</tr>
<tr>
<td>Geothermal</td>
<td>12</td>
<td>Requires 10% of earth’s land surface to be covered with switchgrass</td>
</tr>
<tr>
<td>Biomass</td>
<td>10</td>
<td>Requires construction of a 1-GW(_{\text{peak}}) power plant every 35 h for the next 40 years. Finite uranium supplies imply need for fast breeder or thorium reactors</td>
</tr>
<tr>
<td>Nuclear</td>
<td>10</td>
<td>Requires 0.16% of the earth’s surface to be covered with 10% efficient solar cells. Total solar power reaching the earth’s surface is 120,000 TW</td>
</tr>
<tr>
<td>Solar</td>
<td>&gt;20</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) CO\(_2\) level in May 2011.

\(^2\) Estimates vary between 30 and 80%.
almost certainly be composed of a mixture of these – and other – technologies, with local circumstances (geography, climate, population density) determining the optimal mix for a particular region. Due to the massive efforts involved in implementing any of these options on a Terawatt scale [3, 4], it is more than likely that fossil fuels will continue to play an important role for the next few decades. In principle, the estimated total fossil fuel reserves can sustain a 25–30-TW energy consumption for at least a few more centuries [5]. This, however, requires efficient capture and storage of CO₂ on an enormous scale, using yet unproven technology.

Of the sources shown in Table 1.1, solar energy is the only source that has the potential to meet all our energy needs. To generate 20 TW of power from the sun, the area to be covered with 10% efficient solar cells is about 816,000 km², which corresponds to an area of about 900 × 900 km². While this appears relatively small when projected onto Africa (Fig. 1.1), it should be realized that this is equal to the total surface area of France and Germany combined. Covering such a large area with solar cells presents a daunting task, even when this is undertaken on a (de-centralized) global scale. To illustrate this one would need to produce on average 650 m² of solar cell panels per second, 24/7 for 365 days per year, for the next 40 years in order to reach 20 TW of peak power.

Fig. 1.1 Area of land that needs to be covered with 10% efficient solar cells in order to generate 20 TW of electrical power
The solar cell market is currently growing by 35–40% per year, and is one of today’s fastest growing markets. In 2010, the global production of solar cells was 16.6 GW peak, resulting in a total installed capacity of ~40 GW_peak. While it is difficult to predict how this market will develop over the next few decades, estimates range from 3 to 7 TW of installed PV capacity in 2050. Irrespective of the uncertainties involved, there can be little doubt that solar energy will become an important component of the energy mix in the decades to come.

1.3 From Solar to Fuel

As the contribution of solar energy to the total energy mix increases, it will become difficult for electricity network operators to cope with the intermittent nature of solar power (day/night cycle, clouds). At a certain point, grid-based storage capacities will be exceeded and large-scale energy storage solutions need to be implemented. One of the more attractive possibilities is to store solar energy in the form of a chemical fuel. The energy of a visible-light photon ranges between 1 and 3 eV, or 100–300 kJ/mol, which is more than sufficient for many chemical synthesis routes. Compared to, e.g., batteries and mechanical or gravity-based storage systems such as flywheels and pumped water reservoirs, chemical fuels combine the advantages of high energy storage densities and ease of transportation. Examples of chemical fuels include hydrogen, methane, methanol, gasoline, diesel, etc.

Except for hydrogen, all of these examples require a source of carbon. While CO_2 is an obvious candidate in view of the environmental concerns discussed in Sect. 1.1, capturing CO_2 from the atmosphere comes at a huge entropic cost because of its dilute nature. Fossil fuel-based power plants seem attractive point-sources of highly concentrated CO_2, but the goal of the exercise was to avoid the use of fossil fuels in the first place. A conceptually more attractive route would be to capture the CO_2 emitted by, e.g., cars, and to reuse it by synthesizing fuels with sunlight as the energy source. This would close the CO_2 loop. One of the challenges that would have to be addressed is to minimize the energy penalty involved in capturing the CO_2. Direct photo(electro)chemical conversion of CO_2 to a fuel seems to be even more challenging, as the electrochemical half-reactions for the conversion of CO_2 to, e.g., methanol or methane involve complex six- and eight-electron transfer steps, respectively.

Based on these considerations, the conversion of solar energy into hydrogen appears to be a much more attractive route. Water is a convenient and abundant source of hydrogen, and there is more than enough water available. A back-of-the-envelope calculation shows that ~3.5 \times 10^{13} L of water is needed to store the energy the world uses in 1 year (4.7 \times 10^{20} J) in the form of hydrogen. This corresponds to 0.01% of the annual rain fall, or 0.000002% of the amount of water in the world’s oceans. The water splitting reaction can be written as follows:

$$2\text{H}_2\text{O} + \text{sunlight} \rightleftharpoons 2\text{H}_2 + \text{O}_2 \quad \Delta G = 237 \text{kJ/mol}. \quad (3.1)$$

The reduction half-reaction is an easy two-electron transfer reaction, while four electrons are involved in the oxidation of water to form oxygen. While by no means trivial, it is considerably easier to photo-oxidize water than it is to photoreduce CO₂. Moreover, hydrogen can be readily converted into electricity—and back again—with fuel cells and electrolyzers. This offers the prospect of a future energy infrastructure based on sunlight, hydrogen, and electricity, as illustrated in Fig. 1.2.

One of the main concerns associated with hydrogen is the difficulty in storing it. While hydrogen has a very high gravimetric energy density, the volumetric energy density is rather low (Table 1.2). Solutions can be found in the form of high pressure storage containers (up to 700 bar), liquid cryo-storage, or physisorption at high-surface area metal organic frameworks (MOFs) or in clathrate hydrate cages.

**Table 1.2** Gravimetric and volumetric energy densities of several fuels (at 1 bar)

<table>
<thead>
<tr>
<th>Fuel</th>
<th>Gravimetric (MJ/kg)</th>
<th>Volumetric (MJ/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coal</td>
<td>24</td>
<td>–</td>
</tr>
<tr>
<td>Wood</td>
<td>16</td>
<td>–</td>
</tr>
<tr>
<td>Gasoline (petrol)</td>
<td>44</td>
<td>35</td>
</tr>
<tr>
<td>Diesel</td>
<td>46</td>
<td>37</td>
</tr>
<tr>
<td>Methanol</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>Natural gas</td>
<td>54</td>
<td>0.036</td>
</tr>
<tr>
<td>Hydrogen</td>
<td>143</td>
<td>0.011</td>
</tr>
</tbody>
</table>

The reduction half-reaction is an easy two-electron transfer reaction, while four electrons are involved in the oxidation of water to form oxygen. While by no means trivial, it is considerably easier to photo-oxidize water than it is to photoreduce CO₂. Moreover, hydrogen can be readily converted into electricity—and back again—with fuel cells and electrolyzers. This offers the prospect of a future energy infrastructure based on sunlight, hydrogen, and electricity, as illustrated in Fig. 1.2.
Another solution is to store hydrogen by forming chemical bonds. This can be in the form of metal hydrides, such as MgH₂, LaNi₅H₆, and LiBH₄, or by using hydrogen and CO₂ to make chemical fuels. The latter is a much easier route than the direct photochemical or electrochemical activation of CO₂. For example, CO₂ and hydrogen can be converted into CO via the slightly endothermic reverse water–gas shift reaction:

\[
\text{CO}_2 + \text{H}_2 \rightleftharpoons \text{H}_2\text{O} + \text{CO} \quad \Delta H = +42 \text{ kJ/mol.} \tag{3.2}
\]

The CO and H₂ (syngas) can be separated out with a membrane and converted to liquid hydrocarbon fuels, such as methanol and diesel, using well-established Fischer–Tropsch technology.

### 1.4 Routes to Solar Hydrogen

Many pathways exist for the conversion of water and sunlight into hydrogen:

- Photoelectrochemical water splitting
- Photocatalytic water splitting
- Coupled photovoltaic – electrolysis systems
- Thermochemical conversion
- Photobiological methods
- Molecular artificial photosynthesis
- Plasma-chemical conversion
- Mechano-catalytic, magnetolysis, radiolysis, etc.

Some of these methods are described in more detail in the recent books by Rajeshwar and Grimes [7, 8]. This book focuses exclusively on the first method in the list: photoelectrochemical water splitting with semiconductor photoelectrodes. There are several reasons why this approach is appealing. One of the main advantages is that hydrogen and oxygen are produced at separate electrodes. This avoids serious safety concerns³ and allows easy separation of these gases without having to pay a heavy energy penalty for postseparation. A second advantage is that it can be carried out at room temperature, i.e., there is no need for large-scale solar concentrators that would limit its application to large central facilities in sunny regions of the world. A third advantage is that a photoelectrochemical water splitting device can be constructed entirely from inorganic materials. This offers a degree of chemical robustness and durability that is difficult to achieve for organic or biological systems.

³ The explosion limits of hydrogen are between 18 and 59%, and the flammability limits are between 4 and 74% (in air).
1.5 Benchmark for Photoelectrochemical Water Splitting

It should be noted that the advantages mentioned above for photoelectrochemical (PEC) systems are equally applicable to coupled photovoltaic–electrolysis systems. All the necessary components of such a system (solar cell, electrolyzer, dc–dc converter) are commercially available, and solar-to-hydrogen efficiencies around 8% have already been demonstrated. This approach can therefore be considered as a benchmark for PEC water splitting. The PEC approach offers two potential advantages over PV + electrolysis. The first advantage lies in the fact that commercial electrolyzers require cell voltages of ~1.9 V in order to reach their optimal operating current densities of ~1 A/cm². Since the thermodynamically required potential for water splitting is 1.23 V, this places an upper limit of 65% (1.23/1.9) on the overall energy conversion efficiency [9]. In contrast, the current density at a semiconductor photoelectrode immersed in water is much smaller (10–20 mA/cm² at most) and the required overpotential is therefore substantially lower. The second advantage is that a PEC system can be constructed as a single, monolithic device. This requires fewer packaging components (frame, glass, connections, etc.) and may lead to significantly lower costs. The cost per kg of hydrogen is in fact the key benchmark figure. Estimated costs for hydrogen produced with PV + electrolysis exceed $8/kg, well above the $2–4/kg target set by the US Department of Energy for future hydrogen production pathways. As discussed in more detail in Chaps. 7 and 8, photoelectrochemical water splitting may offer a route toward hydrogen production costs of $3–5/kg, which is competitive with existing energy sources.

1.6 Materials for PEC Devices

The key component for PEC systems is the semiconductor photoelectrode. The ideal photoelectrode fulfills several tasks at once: light absorption, charge separation, charge transport, and H₂ or O₂ evolution at its surface. Moreover, it needs to be stable in an aqueous solution, and have the potential to be made at low cost. No semiconducting material has yet been found that comes even close to meeting these contradictory demands. This means that trade-offs have to be made, leading to the development of composite photoelectrodes in which different materials fulfill different functionalities.

Figure 1.3 illustrates some of the approaches currently being studied in the field. The left-hand figure shows a monolithic PEC device developed by John Turner’s group at NREL [9]. It is based on a p-type GaInP₂ photocathode that is biased by an

---

4 This depends on the catalytic activity of the semiconductor surface, which can be quite low. By attaching suitable co-catalysts, overpotentials as low as 0.3–0.4 V can be achieved.
integrated GaAs pn junction. With an impressive solar-to-hydrogen efficiency of 12.4%, this example is often quoted as the efficiency benchmark for PEC devices. Challenges to be addressed for this approach are the inherently poor stability of the employed semiconductor materials in water (lifetime only a few hours) and the prohibitively high cost of the GaAs and GaInP₂ compounds. The approach shown in the right-hand side of Fig. 1.3 represents a much cheaper solution. It was reported by the group of Thomas Mallouk at Penn State University, and consists of a dye-sensitized nanoporous TiO₂ system in which water oxidation was catalyzed by attaching an IrO₂ nanoparticle to the ruthenium-based dye molecule. The dye molecule and the TiO₂ function as optical absorption centers and charge separators, respectively. The challenges for this system are to improve the performance (currently less than 1% quantum yield) and the degradation of the Ru-based dye by optimizing the kinetic pathways for the reactions of the photogenerated charge carriers.

The middle part of Fig. 1.3 shows an approach in which metal oxide semiconductors, such as Fe₂O₃ and WO₃ are used as light absorbers and charge separators. Inorganic catalysts such as IrO₂ or cobalt-based compounds are generally necessary to catalyze the oxygen evolution reaction. These systems are studied by the authors’ groups at EPFL [10, 12] and TU Delft [13, 14], and many other groups around the world. The main advantage of using metal oxides is their low cost and excellent stability against (photo)corrosion in aqueous solutions – although this does depend on the pH of the solution and the choice of metal oxide. One of the main challenges for metal oxide-based absorbers is the combination of modest light absorption and poor charge transport properties. Solutions for these and other challenges have been proposed in the form of mesoporous materials, guest–host nanostructures, tandem junctions, plasmonics, and combinatorial search methods for new metal oxide semiconductors. Chapters 4, 5, 6, 7, and 9 discuss the exciting developments that have been reported in this field over the past few years.
Acknowledgment  Both authors thank the European Commission’s Framework 7 program (NanoPEC, Project 227179) for support.

References

Chapter 2
Principles of Photoelectrochemical Cells

Roel van de Krol

2.1 The Photoelectrochemical Cell

Figure 2.1 shows a simplified energy diagram of a photoelectrochemical (PEC) cell based on a single photoanode and a metal counter electrode. More complicated configurations that involve photocathodes and/or more than one photodeodelectrode are discussed at the end of this chapter. The main component of the PEC cell is the semiconductor, which converts incident photons to electron–hole pairs. These electrons and holes are spatially separated from each other due to the presence of an electric field inside the semiconductor, the origin of which is discussed in Sect. 2.5. The photogenerated electrons are swept toward the conducting back-contact, and are transported to the metal counter-electrode via an external wire. At the metal, the electrons reduce water to form hydrogen gas. The photogenerated holes are swept toward the semiconductor/electrolyte interface, where they oxidize water to form oxygen gas.

For an alkaline electrolyte, the reduction and oxidation reactions can be written as

\[ 4H_2O + 4e^- \rightleftharpoons 2H_2 + 4OH^- \quad E_{\text{red}}^0 = -0.828 \text{ V vs. NHE} \quad (2.1) \]

\[ 4OH^- + 4h^+ \rightleftharpoons 2H_2O + O_2 \quad E_{\text{ox}}^0 = -0.401 \text{ V vs. NHE} \quad (2.2) \]

Note that the sign of the potential for the oxidation half-reactions is opposite from that usually encountered in the literature, which usually lists these reactions as reduction reactions.

R. van de Krol
Faculty of Applied Sciences, Department Chemical Engineering/Materials for Energy Conversion and Storage, Delft University of Technology, P.O. Box 5045, 2600 GA Delft, The Netherlands
E-mail: r.vandekrol@tudelft.nl

For an acidic environment, the appropriate reactions can be obtained from (2.1) and (2.2) by subtracting or adding the dissociation reaction of water into protons and hydroxyl ions:

\[
4\text{H}^+ + 4\text{e}^- \rightleftharpoons 2\text{H}_2 \quad E^0_{\text{red}} = +0.000 \text{ V vs. NHE} \quad (2.3)
\]

\[
2\text{H}_2\text{O} + 4\text{h}^+ \rightleftharpoons 4\text{H}^+ + \text{O}_2 \quad E^0_{\text{ox}} = -1.229 \text{ V vs. NHE.} \quad (2.4)
\]

The Gibbs free energy change for the overall water splitting reaction is given by the expression:

\[
\Delta G = -nF\Delta E. \quad (2.5)
\]

At standard temperature (298 K) and concentrations (1 mol/L, 1 bar), the electrochemical cell voltage \( \Delta E \) of \(-1.229 \text{ V}\) corresponds to a Gibbs free energy change of +237 kJ/mol \( \text{H}_2 \). This shows that the water-splitting reaction is thermodynamically uphill. This is markedly different from the photocatalysis reactions that one encounters in, e.g., photo-assisted degradation of organic pollutants, for which the Gibbs free energy change is negative.

### 2.2 Semiconducting Photoelectrode Materials

Some of the key requirements for a semiconductor photoelectrode are efficient absorption of visible light and good charge transport. It is often – though not always – easy to determine these parameters from an experiment on a particular material. Clearly, this approach becomes impractical if one wants to screen an entire class of candidate photoelectrode materials. For such cases a more fruitful
approach may be to calculate the electronic structure of a base material, and use this to predict how the properties depend on, e.g., composition. While still far from trivial, electronic structure calculations are now becoming more and more routine. The required computing power and software are readily available, and the number of electronic structure calculations reported in the literature, even by experimentally oriented groups, increases rapidly. However, in order to use these spectra to predict certain photoelectrode properties, one first needs to understand how chemical bonding between the atoms affects the electronic structure. Some of the main principles are discussed below. In contrast to most standard textbooks, we emphasize the properties of metal oxide semiconductors.

In most conventional semiconductors, such as Si and Ge, covalent bonding dominates. In silicon, for example, the outer 3s and 3p orbitals combine to form hybrid sp³ orbitals. Neighboring sp³ orbitals interact to form bonding and antibonding combinations that form the valence and conduction bands of the material, respectively. This is schematically illustrated in Fig. 2.2, which also shows the electronic band structure of silicon.

The bonding in metal oxide semiconductors is very different in nature. Since oxygen has a much higher electronegativity than any metal, the valence electrons are either fully or partially transferred from the oxygen to the metal ion. The bonding character of metal oxides is therefore highly polar or even ionic. A qualitative band picture can be obtained by constructing a molecular orbital (MO) diagram from the individual atomic energy levels. Figure 2.3 shows an example for rutile TiO₂, the very first and most extensively investigated photoanode material for water splitting [2, 3]. The main features of the MO diagram correspond quite well to the calculated band structure for rutile TiO₂, which is shown in Fig. 2.4. The valence band is mainly composed of O-2p orbitals, whereas the conduction band is primarily Ti-3d in character. One could think of the valence band as being occupied with the electrons that originally resided on the titanium

---

**Fig. 2.2** Left: Formation of the valence and conduction bands in covalent semiconductors from bonding and antibonding sp³ orbitals, respectively. Right: Calculated electronic band structure of silicon [1]. The gray area indicates occupied states in the valence band of the material.
Fig. 2.3 Molecular orbital diagram of rutile TiO$_2$ (after Stoyanov [4] and Fisher [5])

Fig. 2.4 Electronic band structure and density-of-states (DOS) of rutile TiO$_2$. The black parts of the DOS indicate completely filled bands. Adapted from Hoffmann [6], pp. 31, copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission
atoms, before they were transferred to the more electronegative oxygen during the formation of the bond.\textsuperscript{2}

It should be noted that such a “local” view is entirely inappropriate for covalent semiconductors such as silicon, in which the electrons are completely delocalized over the material. This local vs. delocalized view is a key feature that distinguishes metal oxide semiconductors from their covalently bonded counterparts. The local character of the energy bands in metal oxides sometimes leads to ambiguous interpretations. For example, “free” electrons in TiO\textsubscript{2} are sometimes distinguished from Ti\textsuperscript{3+} species. While such a distinction may be realistic at the surface of a material, where the electronic structure is distorted due to a disruption of the lattice symmetry, it is not obvious in the bulk in which the conduction band is mainly composed of Ti 3d orbitals.

While seemingly complicated, the electronic band structure can directly lead to several important insights for photoelectrode materials. The first insight is the nature of the optical transition, illustrated in more detail in Fig. 2.5. If the highest point in the valence band is located at the same $k$-vector as the lowest point of the conduction band, the optical transition does not require a change in crystal momentum. Such a transition is called direct, and is indicated by a vertical line in the $E$–$k$ diagram. In contrast, indirect transitions do require a change in crystal momentum, as illustrated in Fig. 2.5. Since photons carry very little momentum, indirect transitions require absorption or emission of a phonon (i.e., a lattice vibration) and are much less likely to occur. As a consequence, the absorption coefficient of indirect semiconductors is much smaller than that of direct semiconductors – the difference can be two orders of magnitude. This is why direct semiconductors, such as amorphous Si or CuIn\textsubscript{x}Ga\textsubscript{1-x}Se\textsubscript{2} (CIGS) can absorb all incident light in just a few micrometers, whereas crystalline (indirect) Si solar cells typically have a thickness

\textsuperscript{2}The valence band of a semiconductor is analogous to the highest occupied molecular orbital (HOMO) in a molecule, whereas the conduction band is the solid state analogue of the lowest unoccupied molecular orbital (LUMO).
of ~300 μm. Metal oxides often have fairly horizontal-running $E$-$k$ lines in their band structure (cf. Fig. 2.4), which sometimes obscures the direct vs. indirect nature.

The bandgap of a material can be determined from a measurement of the absorption coefficient vs. wavelength. If the bottom of the conduction band and the top of the valence band are assumed to have a parabolic shape, the absorption coefficient (units: m$^{-1}$) can be expressed as follows:

$$\alpha = A \frac{(\hbar \nu - E_g)^m}{\hbar \nu}.$$  \hspace{1cm} (2.6)

Here, $A$ is a constant and $m$ depends on the nature of the optical transition: $m = \frac{1}{2}$ for a direct bandgap, and $m = 2$ for an indirect gap. From (2.6), extrapolation of a plot of $(\alpha \hbar \nu)^{1/2}$ vs. $\hbar \nu$ plot gives the indirect bandgap, while a plot of $(\alpha \hbar \nu)^2$ vs. $\hbar \nu$ yields the direct bandgap of the material. Such a plot is called a “Tauc plot” [7] and is often encountered in the photoelectrochemistry and photocatalysis literature.

The second insight that can be obtained from the electronic band structure is the mobility of charge carriers, which is related to the width of the conduction and valence bands. For Si the bands are rather broad, spanning more than 10 eV. This is a direct consequence of the extensive overlap of the sp$^3$ orbitals on neighboring atoms. More overlap between atomic wavefunctions results in broader bands and easier transport of free charge carriers through the material. This can be quantified via the curvature of the individual bands, which is directly related to the effective mass and mobility of the charge carriers:

$$m^* = \frac{h^2}{(\frac{dE}{dk})^2} \quad \text{and} \quad \mu = \frac{q \tau}{m^*}. \hspace{1cm} (2.7)$$

Wide bands that are strongly curved indicate a high charge carrier mobility. A high mobility is especially desirable in photoelectrodes with an indirect bandgap. This is because these materials require a large thickness to absorb all the incident light, which means that photogenerated electrons and holes have to travel large distances before reaching the interface.

Table 2.1 gives an overview of the room temperature charge carrier mobilities of several semiconductors. What stands out are the small charge carrier mobilities for the transition metal oxides. One reason for this is the fact that the 3d orbitals in transition metals are fairly contracted, and show less overlap with neighboring atoms than the s- and p-orbitals. One can also look at this from an electrostatic point of view. The electrostatic interaction of a free electron with the cations surrounding it creates a local lattice distortion. This distortion accompanies the electron when it moves through the lattice, and is called a polaron. If the interaction is very strong, the polaron can even become trapped at a particular lattice site. The electron can then only move by thermally activated hopping, and its mobility will be very low. Similar considerations apply to “free” holes in the valence band, which can form polarons through electrostatic interaction with the surrounding oxygen ions. Charge transport via polarons is believed to play an important role in, e.g., hematite ($\alpha$-Fe$_2$O$_3$), which will be discussed in more detail in Chap. 4.
The third piece of information that can be obtained from the electronic band structure is the density-of-states (DOS). This is the number of allowed electronic states per unit of energy interval. In $E-k$ diagrams the $k$-values are equally spaced along the $k$-axis, so the DOS can be readily estimated from the total length of the black lines that run through a certain interval $D_E$. As illustrated in Fig. 2.4 for TiO$_2$, the highest DOS occur at energies where the $E-k$ curves are flat (horizontal). The DOS representation not only allows one to quickly and easily identify the various bands and band widths, but it also gives a rough impression of the optical transition probabilities between various bands. This is because the total number of excitations per second is proportional to the DOS at the occupied ground state (the initial state), multiplied by the transition probability to the final state, $\lambda_{if}$. The latter is given by Fermi’s Golden Rule, (2.8), and depends linearly on the DOS at the final (excited) state, $g_f$:

$$\lambda_{if} = \frac{2\pi}{\hbar} |M_{if}|^2 g_f. \tag{2.8}$$

In some cases, one can actually see certain features of the DOS reflected in the optical absorption spectrum of the material.

### Table 2.1 Charge carrier mobilities and effective masses at room temperature

<table>
<thead>
<tr>
<th>Material</th>
<th>Electron Mobility (cm$^2$/V/s)</th>
<th>Hole Mobility (cm$^2$/V/s)</th>
<th>Effective Electrons Mass ($m_e^*$)</th>
<th>Effective Holes Mass ($m_h^*$)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>1,500</td>
<td>450</td>
<td>0.07</td>
<td>0.68</td>
<td>[8]</td>
</tr>
<tr>
<td>GaAs</td>
<td>8,500</td>
<td>400</td>
<td>0.19</td>
<td>0.8</td>
<td>[8, 9]</td>
</tr>
<tr>
<td>CdS</td>
<td>340</td>
<td>50</td>
<td>0.24</td>
<td>0.45</td>
<td>[8, 9]</td>
</tr>
<tr>
<td>ZnO</td>
<td>200</td>
<td>180</td>
<td>$2 \times 10^{-3}$ a</td>
<td>&gt;10</td>
<td>[10, 11]</td>
</tr>
<tr>
<td>TiO$_2$ (anatase)</td>
<td>2 × 10$^{-3}$ a</td>
<td>&gt;10</td>
<td>0.8</td>
<td>[12]</td>
<td></td>
</tr>
<tr>
<td>WO$_3$</td>
<td>~10</td>
<td>~0.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$-Fe$_2$O$_3$</td>
<td>0.1$^b$</td>
<td>0.2$^c$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-Cu$_2$O</td>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-CuAl$_2$O$_2$</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-SrCu$_2$O$_2$</td>
<td>0.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-CuMn$_2$O$_2$</td>
<td>3.5 × 10$^{-6}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$Theoretical estimate
$^b$At 1,200 K
$^c$At 1,400 K

2.3 Charge Carriers and Doping

Under equilibrium conditions (i.e., no illumination and no net current flow), the concentration of free electrons in the conduction band and free holes in the valence band is given by the following expressions:

$$n = N_C e^{-(E_C - E_f)/kT} \quad \text{with} \quad N_C = 2 \left( \frac{2\pi m^*_e kT}{\hbar^2} \right)^{3/2}, \tag{2.9}$$
\[ p = N_V e^{-(E_F-E_V)/kT} \quad \text{with} \quad N_V = 2 \left( \frac{2\pi m^*_h kT}{\hbar^2} \right)^{3/2}. \]  

Equations (2.9) and (2.10) are valid for both undoped and doped semiconductors. They are, however, not valid when the Fermi level is less than \( \sim 3kT \) away from either one of the band edges. Under these conditions, the semiconductor is degenerate, and exhibits near-metallic behavior. The relationships for the effective densities of states were derived from the (nearly) free electron model, and may not be entirely accurate for transition metal oxides. Despite these limitations, (2.9) and (2.10) are exceedingly useful for describing the behavior of semiconducting photoelectrodes.

Semiconductors for practical applications are often doped, mainly with the aim to improve the conductivity. In metal oxide photoelectrodes, shallow donors and acceptors are almost always necessary because of the low intrinsic charge carrier mobilities. The conductivity of the material is given by \( \sigma = ne\mu_e + p\mu_h \), so increasing \( n \) or \( p \) will compensate for a small value of \( \mu_e \) or \( \mu_h \). Examples of donor-type dopants are Ti\(^{4+}\) on a Fe\(^{3+}\) site in Fe\(_2\)O\(_3\), or phosphorus in silicon. The extra valence electron introduced by the donor atom is loosely bound to the donor nucleus, and can be excited to the conduction band where it then contributes to the conductivity. Conversely, holes in acceptor-type dopants can be excited to the valence band. Since a hole is equivalent to a missing electron, one can also picture this as an electron being excited from the valence band into the energy level of the acceptor species.

If the dopant level is within \( \sim 2kT \) of \( E_C \) or \( E_V \), it will be (almost) fully ionized at room temperature – this is referred to as a shallow dopant. For deep donors and acceptors, the degree of ionization can be calculated with the following equations:

\[ \frac{N^+_D}{N_D} = \frac{1}{1 + g_D e^{(E_f-E_D)/kT}} \quad \text{and} \quad \frac{N^-_A}{N_A} = \frac{1}{1 + g_A e^{(E_A-E_f)/kT}}. \]  

Here, \( E_D \) and \( E_A \) are the donor and acceptor energies and \( g_A \) and \( g_D \) are the corresponding degeneracy factors that reflect the multiplicity of the energy state [18]. Under the assumption that \( n = N^+_D \), the following useful expressions relate the free electron concentration directly to the position of the donor level in the bandgap [18]:

\[ n = \frac{N_z}{2} \left( \sqrt{1 + \frac{4N_D}{N_z}} - 1 \right) \quad \text{with} \quad N_z = \left( \frac{N_c}{g_D} \right) e^{-(E_C-E_D)/kT} \]  

(2.12)
An analogous expression can be obtained for acceptor-doped materials. The energy levels for the various dopants are illustrated in Fig. 2.6. Deep dopants can act as optically active centers or as catalytically active surface sites, as given in the following paragraphs.

2.4 Defect Chemistry

Semiconducting photoelectrodes are almost always doped to improve their properties. In most cases, the aim is to enhance the n- or p-type conductivity, as described in Sect. 2.3. Certain dopants may enhance the optical absorption of wide bandgap semiconductors [19], increase the minority carrier diffusion length [20, 21], or enhance the catalytic activity at the surface of the semiconductor [22]. Other dopants adversely affect the properties, for example, by introducing midgap bulk or surface states that act as recombination centers [23, 24].

In addition to foreign cation and anion dopants, native point defects are also usually present in the material. Examples are vacancies, interstitials, or substituents. These defects can have a similar influence on the optical, electrical, and catalytic properties as dopants. They are formed by intrinsic defect-chemical reactions, or by a change in the lattice stoichiometry due to exchange of, e.g., oxygen with the gas phase. Since virtually every defect affects the performance of the material in some way, the ability to understand and predict the relationship between dopants and defect concentrations is of paramount importance for designing efficient photoelectrodes.

---

3 An example of a native substituent is a site exchange of A and B cations in a ternary compound such as ABₓOᵧ.
2.4.1 Doping Binary Metal Oxides

In covalently bonded semiconductors, the introduction of dopants is fairly straightforward since all lattice sites are similar to each other. This is even true for a polar compound such as GaAs. In this material, Si can be incorporated on both Ga and As lattice sites [25]. For ionically bonded materials the situation is quite different, since the large electrostatic penalty prevents any exchange of anions and cations. This puts certain restrictions on the incorporation of dopants, as given below.

The incorporation of dopants in metal oxides and other ionic materials can be conveniently described with the so-called Kröger–Vink notation [26], which is summarized in Fig. 2.7. For a more extensive description the reader is referred to several excellent textbooks on this subject [27, 28]. When describing defect-chemical reactions, one has to observe the conservation of mass and charge, just as one would for any other chemical (half)reaction. What is different in the defect chemistry of ionic solids is the conservation of lattice site stoichiometry. For example, if we dope Fe$_2$O$_3$ with zinc oxide, two zinc ions substitute for Fe, and the two oxygen ions occupy two of the three oxygen sites in the Fe$_2$O$_3$ unit. This means that the third oxygen site remains empty, resulting in an oxygen vacancy:

$$2\text{ZnO} \xrightarrow{\text{Fe}_2\text{O}_3} 2\text{Zn}^\text{Fe} + 2\text{O}_2^\text{O} + \text{V}_\text{O}^{**}.$$  (2.13)
Note that this is an irreversible reaction, since spontaneous demixing will not occur. This means that one cannot define an equilibrium constant for a dissolution reaction. Closer inspection shows that reaction (2.13) indeed fulfills the required conservation of mass, charge, and lattice site stoichiometry.

One may try to get rid of the oxygen vacancy in reaction (2.13) by performing the synthesis in an oxygen-rich atmosphere. In this case, both mass and lattice site stoichiometry are conserved when adding the oxygen gas, which means that the charge has to be balanced by the addition of a hole:

$$2\text{ZnO} + 1/2\text{O}_2(g) \xrightarrow{\text{Fe}_2\text{O}_3} 2\text{Zn}^{+\text{Fe}} + 3\text{O}_\text{O}^{\text{O}} + h^\bullet.$$  \hspace{1cm} (2.14)

Note that the presence of a hole suggests p-type conductivity in Fe$_2$O$_3$. In many cases, however, metal oxides remain n-type when acceptor-doped. If this is the case, it is more appropriate to replace the hole with an electron on the left-hand side of (2.14).

In (2.13) and (2.14) the charge of the aliovalent$^4$ Zn dopant is compensated by an ionic defect ($V_\text{O}^{\text{O}}$) and an electronic defect ($h^\bullet$), respectively. To illustrate the difference between these ionic and electronic compensation mechanisms in more detail, consider Ti-doped Fe$_2$O$_3$. When subtracting the ionic compensation reaction from the electronic one, we obtain

$$6\text{TiO}_2 \rightarrow 3\text{Fe}_2\text{O}_3 6\text{Ti}^{\text{Fe}} + 9\text{O}_\text{O}^{\text{O}} + 3/2\text{O}_2(g) + 6e^\bullet$$ \hspace{1cm} (2.15)

$$6\text{TiO}_2 \rightarrow 4\text{Fe}_2\text{O}_3 6\text{Ti}^{\text{Fe}} + 12\text{O}_\text{O}^{\text{O}} + 2V^{\text{Fe}}$$ \hspace{1cm} (2.16)

$$3\text{O}_\text{O}^{\text{O}} + 2V^{\text{Fe}} \Leftrightarrow 3/2\text{O}_2(g) + 6e^\bullet.$$  \hspace{1cm} (2.17)

This shows that the difference between two irreversible dissolution reactions can be written as a reversible oxidation/reduction reaction. An important implication of this is that the addition of a dopant is, by itself, not sufficient to enhance the conductivity of a metal oxide photoelectrode: one also needs to ensure that the equilibrium of reaction (2.17) lies at the right-hand side. The factors that affect the equilibrium position are oxygen partial pressure, dopant concentration, and temperature. The equilibrium constant of (2.17) is given by

$$K = \frac{n^6P^{3/2}}{[V_{\text{Fe}}^{\text{Fe}}]^2}.$$  \hspace{1cm} (2.18)

$^4$An “aliovalent” dopant has different charge than the ion that it replaces.
From this, it is immediately clear that a low $p(O_2)$ favors electronic compensation. Moreover, $K$ is constant, so at constant $p(O_2)$ the factor $n^6/[V_{Fe}^{3+}]^2$ also remains constant. This implies that at high dopant concentrations, ionic compensation will dominate over electronic compensation. In other words, a plot of the conductivity as a function of the Ti concentration will have its highest slope at low dopant concentrations. Finally, we note that the entropy ($\Delta S$) of (2.17) strongly increases from left to right. Since $\Delta G = \Delta H - T\Delta S$, electronic compensation is favored at high temperatures.

It is interesting to note that after subtracting the Fe$_2$O$_3$ Schottky reaction (see Fig. 2.7) from (2.17), the standard oxygen reduction reaction for n-type oxides is obtained:

$$\text{O}_2^* \rightleftharpoons V_{O}^{3+} + 1/2 \text{O}_2(g) + 2e^-.$$ (2.19)

This reaction is of course valid for both doped and undoped metal oxides.

For doped metal oxides, the aliovalent metal ions influence the overall defect concentrations via what is sometimes called the “First Law of Doping.” This law simply states that adding an aliovalent dopant increases the concentration of defects with opposite charges, and decreases the concentration of defects with charges of the same sign.

The concentration of dopants in metal oxides is usually limited to 1–2% at most, which corresponds to a concentration of $\sim 10^{21}$ cm$^{-3}$. Higher concentrations are not likely to be effective, and may even lead to segregation of the dopant phase. If the conductivity of a doped metal oxide photoelectrode is found to be too low, it may be more effective to anneal it at high temperatures and under low $p(O_2)$ than to increase the dopant concentration.

In (2.13)–(2.16), only one type of ionic defect was considered for each ionic compensation reaction. In the case of ZnO-doped Fe$_2$O$_3$, the Zn acceptor could also be compensated by a cation interstitial (Fe$^{3+}$) instead of an oxygen vacancy. However, the higher charge of the interstitial Fe makes this energetically less favorable. Similarly, oxygen interstitials instead of iron vacancies could compensate the Ti donor in reaction (2.16), but in this case the large size of interstitial anions prohibits this. In general, donor dopants are compensated by cation vacancies or electrons, while acceptor dopants are compensated by anion vacancies, cation interstitials, or holes. Which ionic defect is more likely to occur depends on the crystal structure, the ionic radius, and the charge of the ion.

---

5 Anion interstitials normally only occur in oxides with the fluorite structure, which can be viewed as an fcc base lattice of cations in which the interstitial sites are occupied by anions.
2.4.2 Dopant Energy Levels

Most defects in metal oxides can have multiple oxidation states. For example, neutral oxygen vacancies can be ionized to divalent species in two steps:

\[
\begin{align*}
V_x^O & \rightarrow V_x^0 + e^- \quad (2.20) \\
V_x^0 & \rightarrow V_x^{0^+} + e^- \quad (2.21)
\end{align*}
\]

The corresponding ionization energies are depicted in the energy diagram in the left-hand panel of Fig. 2.8. The common convention is to label the donor states by their charge before ionization, whereas acceptor states are labeled by their charge after ionization. Hence, D_Mx represents a shallow donor, and A_M a deep acceptor.

As outlined above, shallow donors are used to increase the conductivity of an n-type oxide. Examples are Ti and Si in Fe₂O₃ [29–31], W in BiVO₄ [32], and Nb in TiO₂ [33]. Oxygen vacancies are often considered to be shallow donors as well, based on the observation that the conductivity of the oxide increases when reducing it at low p(O₂). However, in many oxides – especially those with a fairly wide bandgap – the ionization energy of a singly ionized oxygen vacancy is simply too high (>3 kT) to be considered as a shallow donor at room temperature. This implies that the conductivity is mostly determined by ionization of neutral oxygen vacancies, i.e., reaction (2.20). Note that this is somewhat different from the mechanism implied by reaction (2.19), in which reactions (2.20) and (2.21) are both assumed to occur. When present as deep
donor species, ionized oxygen vacancies may adversely affect the performance of photoanodes by acting as recombination centers. If this is the case, their concentration can be kept low by doping with shallow donors, i.e., by using the First Law of Doping.

Cr and Fe are well-known examples of deep acceptors in TiO₂. As indicated in Fig. 2.8 (right), the electron in these acceptor levels can be optically excited to the conduction band. This requires less energy than the bandgap of the material, and Fe- and Cr-doped TiO₂ actually absorb visible light. An example of this is shown in Fig. 2.9 (left) for Cr-doped TiO₂. This visible light sensitization effect has attracted much interest in the field of photoelectrochemistry, especially in the late 1970s and early 1980s. A problem with this approach is that hole transport toward the interface is slow, as it takes place via hopping along neighboring acceptor levels (cf. Fig. 2.8). Moreover, the optical absorption cross section of these defects is small, requiring relatively thick films and long hole diffusion path lengths. Because of this, most holes recombine before reaching the interface. The photocurrents for Fe- or Cr-doped TiO₂ are indeed found to be low, and use of cation doping to improve the visible light absorption of wide-bandgap semiconductors has all but been abandoned in photoelectrochemistry. However, it still attracts interest in photocatalysis based on doped TiO₂ nanoparticles, where the holes only have to travel a short distance and recombination is less severe.

Interestingly, TiO₂ doped with Cr by physical ion implantation shows less recombination, and the shape of the absorption spectrum suggests an actual decrease of the bandgap instead of a local absorption center [34]. An example is shown in the right-hand side of Fig. 2.9. Clearly, the different incorporation methods for Cr give rise to different defects and a different electronic structure of the material. The physical reason for these differences remains to be established.

Fig. 2.9  Left: Optical absorption spectrum of TiO₂ doped with Cr via chemical impregnation. The amount of Cr increases from 0 wt% (a) to 1 wt% (e'). Right: Optical absorption of TiO₂ doped with Cr by physical implantation, with the Cr concentration increasing from 0 wt% (a) to 1.3 wt% (d). Reprinted from [34], with permission from Elsevier
2.4.3 Defect Equilibria

If all the defect formation energies and/or equilibrium constants are known, the concentration of all the defects in a material can be calculated. As an example, we briefly consider the case for a hypothetical M$_2$O$_3$ photoanode doped with X$^{4+}$. Let us assume that X substitutes for the metal (as opposed to occupying an interstitial site), that Schottky disorder dominates in M$_2$O$_3$, and that the material is n-type ($n \gg p$). We furthermore assume that fully ionized oxygen vacancies are deep donors and can easily trap an electron, whereas singly ionized oxygen vacancies are shallow donors. We can then write the following defect equations and equilibrium constants:

\[
\begin{align*}
2XO_2 & \rightarrow 2X_M^\bullet + 3O_6^O + 1/2O_2(g) + 2e^- \\
0 & \rightleftharpoons 3V^{///}_M + 2V^{***}_O \quad K_S = [V^{///}_M]^3[V^{***}_O]^2 \\
0 & \rightleftharpoons e^\prime + h^\bullet \quad K_e = np \\
O_6^O & \rightleftharpoons V^{***}_O + 1/2O_2(g) + 2e^- \quad K_R = n^2[V^{***}_O]^1/2 \\
V^{\bullet}_O & \rightleftharpoons V^{***}_O + e^\prime \quad K_i = \frac{[V^{***}_O]^n}{[V^{\bullet}_O]}.
\end{align*}
\]

No equilibrium constant is defined for the dissolution reaction, since it is irreversible. Note that we have chosen to write the electronic compensation form of the dissolution reaction; we could have just as easily have chosen the ionic compensation reaction. As we saw before, they are related through reduction reaction (2.25) and are therefore not independent of each other.

We now have four equilibrium equations and five unknown variables ($[V^{///}_M], [V^{***}_O], [V^{\bullet}_O], n, p$). The missing equation is the electroneutrality equation:

\[
3[V^{///}_M] + n = 2[V^{***}_O] + [V^{\bullet}_O] + p \quad (2.27)
\]

This system of nonlinear algebraic equations can be numerically solved with a computer, or one can find analytical solutions by simplification of (2.27) into six separate Brouwer regimes. Each Brouwer regime is characterized by a single dominant positive defect and a single dominant negative defect. While finding the right order of the different Brouwer regimes as a function of $p(O_2)$ may pose a bit of a challenge, one can derive all defect concentrations as a function of $p(O_2)$ by hand and plot these in a so-called “Brouwer diagram” (sometimes also...
2.4.4 Doping of Complex Metal Oxides

Since only very few binary oxides show promise as photoelectrodes for water splitting, most notably WO₃ and Fe₂O₃, PEC research activities have now partly shifted toward ternary and more complex metal oxides (see, e.g., Chaps. 5 and 6). One of the best known examples is BiVO₄, which is an n-type semiconductor with a
bandgap of ~2.4 eV [35, 36]. Recent results from the author’s laboratory show that electron transport limits the performance of undoped BiVO₄ photoanodes [32]. To improve this, one can dope BiVO₄ with tungsten in order to increase the concentration of free electrons. Tungsten has a valence of 6+ and a crystal ionic radius of 42 pm. It will therefore be a substitute for tetrahedrally coordinated V⁵⁺, which has a slightly smaller ionic radius (36 pm).⁶ When W⁶⁺ substitutes for V⁵⁺, one might expect the positive charge to be compensated by free electrons, thereby increasing the conductivity. However, there is a minor complication: the compensating negative charge is already provided in the form of Bi vacancies. Since tungsten only occupies sites on the vanadium sublattice, the formation of Bi vacancies is necessary in order to preserve lattice site stoichiometry. This leads to the following electronic compensation reaction:

\[
\text{WO}_3 + 1/2 \text{O}_2 (g) + 2e^- \xrightarrow{\text{BiVO}_4} \text{V}^{\text{III}}_{\text{Bi}} + \text{W}^\bullet + 4\text{O}^\bullet_{\text{O}}.
\]  

(2.28a)

According to (2.28a), W-doping actually decreases the conductivity of n-type BiVO₄, which is a rather unexpected (and undesired) result. It is also possible to write an alternative reaction in which lattice site stoichiometry is preserved by the segregation of vanadium out of the BiVO₄ in the form of V₂O₅:

\[
2\text{BiVO}_4 + 2\text{WO}_3 \xrightarrow{2\text{BiVO}_4} 2\text{Bi}^\bullet_{\text{Bi}} + 2\text{W}^\bullet + \text{V}_2\text{O}_3 + \text{O}^\bullet_{\text{O}} + 1/2 \text{O}_2 (g) + 2e^-.
\]  

(2.28b)

This avoids the need for highly charged cation vacancies, and the W dopant is now indeed charge-compensated by the desired free electrons. Which reaction actually occurs – (2.28a) or (2.28b) – depends on the energy required to form a triply charged Bi vacancy vs. the energy required to form a second phase.

It should be realized that reactions (2.28a) and (2.28b) are both unfavorable from an energetic point of view. In one case a highly charged defect has to be formed, while in the other case a second phase has to segregate out. This leads to an important general observation for ternary – and more complex – metal oxides: doping these materials by simply dissolving a binary oxide is difficult to achieve in practice. An elegant solution for this is to co-dope the BiVO₄ with equal amounts of W and Bi:

\[
2\text{WO}_3 + \text{Bi}_2\text{O}_3 \xrightarrow{2\text{BiVO}_4} 2\text{Bi}^\bullet_{\text{Bi}} + 2\text{W}^\bullet + 8\text{O}^\bullet_{\text{O}} + 1/2 \text{O}_2 (g) + 2e^-.
\]  

(2.29)

Co-doping can be easily achieved by having a small excess of Bi over V during the synthesis of the material, so that \([\text{Bi}] + [\text{W}] = [\text{V}]\). This is a well known and powerful trick to increase the solubility of binary oxides in ternary oxides in order to incorporate donor- or acceptor-type dopants.

⁶Note that the octahedrally coordinated Bi sites (103 pm) are clearly too large for W, and would cause the W⁶⁺ to “rattle,” which is energetically very unfavorable.
2.5 Space Charges and Band Bending

One of the key features of a semiconductor is the presence of a built-in electric field, or space charge. In devices for solar energy conversion, this field is responsible for the efficient separation of the photogenerated electrons and holes, thereby preventing recombination. The field is present near the surface or interface of a semiconductor, and is formed by the transfer of charges from the semiconductor bulk across its interface. The reason for charge transfer and the field induced by the remaining charges inside the semiconductor is described below.

2.5.1 Origin of the Space Charge Layer

In most semiconductor textbooks, the formation of a space charge is explained by a semiconductor being brought into contact with a metal or another semiconductor [8, 37]. When both materials have different Fermi levels (i.e., electrochemical potentials), charge is transferred between them until an equilibrium is established. In metal oxide semiconductors, a built-in electric field can even be formed at the surface without making contact with another solid. This is illustrated in the schematic diagram of Fig. 2.11. When a metal oxide is exposed to air, water molecules from the air can dissociatively adsorb at its surface, resulting in -OH surface termination. Since the symmetry of the bulk lattice is broken at the surface, these -OH groups form electronic surface states within the bandgap of the semiconductor. The energy levels of these surface states are below the conduction band minimum, and free electrons from the bulk will occupy these levels. The ionized donors from which the free electrons originated will of course stay behind in the bulk, where they form a positive space charge. An electric field now forms, and the charge transfer from bulk to surface will continue until the potential barrier becomes too large for bulk electrons to cross. At this point, a dynamic equilibrium establishes at which no net electron transport takes place. The Fermi level at the surface is then located somewhere halfway the DOS of the surface state.

2.5.2 Potential Distribution in The Depletion Layer

The potential distribution and width of space charge depend on the amount of charges transferred to the surface and the density of shallow donors in the material, \( N_D \). Quantitative expressions for these relationships are indispensable in PEC research, and to understand where they come from, we derive them below. The geometry of the problem and the relevant parameters are schematically illustrated in Fig. 2.12. The starting point for the derivation is Poisson’s law, which relates the potential to the net amount of charge:

\[
\frac{d^2 \phi}{dx^2} = \frac{-d \xi}{dx} = \frac{-\rho(x)}{\varepsilon_0 \varepsilon_r}.
\]  

\[ (2.30) \]
Fig. 2.11  Simplified illustration of the formation of a space charge region (SCR) at a metal oxide semiconductor surface when exposed to (humid) air

Fig. 2.12  Illustration of band bending at the surface of a n-type semiconductor in depletion. The “+” markers just below the conduction band represent the ionized donor species, whereas the “−” markers just above the conduction band level represent the free electrons. The potential in the bulk of the semiconductor is chosen as zero
Here, $\zeta$ was chosen as a symbol for the electric field to avoid confusion with symbol $E$ for energy. The total charge density in the space charge is given by the concentration of free electrons, $n$, plus the concentration of ionized donors, $N_D^+$:

$$\rho(x) = e(N_D^+ - n). \quad (2.31)$$

The concentration of free electrons is given by the distance between the conduction band minimum and the Fermi level, cf. (2.9). Inside the SCR, this distance is modified by the band bending $f(x)$, while leading to the following expression for $n$:

$$n(x) = N_C e^{-(E_C - E_F - e\phi(x))/kT} = n_b e^{e\phi(x)/kT}. \quad (2.32)$$

Note that the energy of an electron is related to its potential by $E = -e\phi$ and that the potential in the bulk of the semiconductor is conveniently chosen as zero (see Fig. 2.12). Since the bulk is electrically neutral, its electron concentration, $n_b$, must be equal to the number of ionized donors. The latter are assumed to be fully ionized, so that $N_D^+ = N_D = n_b$. Combining this with (2.31) and (2.32), the total charge density at point $x$ can be written as

$$\rho(x) = eN_D \left( 1 - e^{e\phi(x)/kT} \right). \quad (2.33)$$

The combination of (2.30) and (2.33) cannot be readily solved due to the implicit dependence on $f(x)$. Fortunately, there is an elegant way around this by taking the derivative of the square of the electric field [38]. Using this with (2.30) yields:

$$\frac{d(\zeta^2)}{dx} = 2\zeta \frac{d\zeta}{dx} = 2\zeta \frac{\rho(x)}{\varepsilon_0\varepsilon_r} = -2 \frac{\rho(x)}{\varepsilon_0\varepsilon_r} \frac{d\phi}{dx}. \quad (2.34)$$

Since there is a one-to-one correspondence between $\phi$ and $x$, we can now switch the independent variable from $x$ to $\phi$ and rewrite (2.34) as

$$\zeta^2 = \int_0^{\phi(x)} -2 \frac{\rho(\phi)}{\varepsilon_0\varepsilon_r} d\phi. \quad (2.35)$$

The total net charge between the (neutral) bulk and point $x$ in the space charge can be related to the electric field at point $x$ via the integral form of Gauss’ law:

$$\zeta = \frac{Q}{\varepsilon_0\varepsilon_r A}. \quad (2.36)$$
Here, $A$ is the surface area of semiconductor. Combining (2.35) and (2.36) gives:

$$Q = \sqrt{-2\varepsilon_0\varepsilon_r A^2 \int_0^\phi \rho(\phi) \, d\phi}. \quad (2.37)$$

This equation provides a convenient starting point for further analysis of the space charge. It is generally applicable and can also deal with nonhomogeneous doping profiles and deep donors/acceptors. Here, we limit ourselves to homogeneously distributed shallow donors, for which the charge density is given by (2.33).

The total amount of charge in the SCR can now be obtained by combining (2.33) and (2.37) and integrating between $\phi = 0$ and $\phi = \phi_S = -\phi_{SC}$:

$$Q_{SC} = \sqrt{-2\varepsilon_0\varepsilon_r A^2 \int_0^{-\phi_{SC}} e N_D (1 - e^{\phi/kT}) \, d\phi}$$

$$= \sqrt{2\varepsilon_0\varepsilon_r e N_D A^2 \left(\phi_{SC} + \frac{kT}{e} e^{-e\phi_{SC}/kT} - \frac{kT}{e}\right)}. \quad (2.38)$$

Under normal PEC operating conditions, the potential drop across the space charge is at least 0.1 V, and the previous result can be simplified to

$$Q_{SC} = \sqrt{2\varepsilon_0\varepsilon_r e N_D A^2 \left(\phi_{SC} - \frac{kT}{e}\right)}. \quad (2.39)$$

The total amount of charge is related to the depletion layer thickness via

$$Q_{SC} = e N_D A W \quad (2.40)$$

which leads to the following widely used expression for the space charge width:

$$W = \sqrt{\frac{2\varepsilon_0\varepsilon_r}{e N_D} \left(\phi_{SC} - \frac{kT}{e}\right)}. \quad (2.41)$$

Although this important result was derived for n-type semiconductors, it is also valid for p-type materials if $N_D$ is replaced with the shallow acceptor density, $N_A$.

To get an idea of typical numbers, Fig. 2.13 shows the calculated depletion layer width for $\alpha$-Fe$_2$O$_3$ as a function of $\phi_{SC}$ for various dopant densities. Typical values range between 5 and 500 nm. The total amount of charge in the depletion layer has to be compensated by a sheet of oppositely charged species (e.g., trapped electrons) at the surface of the material. As can be seen in the right-hand part of Fig. 2.13, less
than 1% of a monolayer\(^7\) of adsorbed surface charges is already more than sufficient to induce a significant space charge layer in the material.

The actual values of the dopant density and the depletion layer width can be determined by impedance measurements. Under certain conditions one can extract the differential capacitance of the space charge, \(C_{SC}\), from the total impedance of the system. \(C_{SC}\) can be derived by differentiating (2.39) with respect to \(\phi_{SC}\):

\[
\frac{1}{C_{SC}^2} = \left( \frac{dQ_{SC}}{d\phi_{SC}} \right)^{-2} = \frac{2}{\varepsilon_0 \varepsilon_r \epsilon N_D A^2 / kT} \left( \phi_{SC} - \frac{kT}{e} \right),
\]

This is the so-called “Mott–Schottky” equation. Typical values for \(C_{SC}\) are 10–1,000 nF/cm\(^2\). By plotting \(1/C_{SC}^2\) as a function of the applied potential, the donor density of the semiconductor can be determined.

### 2.5.3 Deep Depletion, Inversion, and Accumulation Layers

In addition to the depletion layer, other types of space charges can also form in a semiconductor. These are illustrated in Fig. 2.14. If the number of adsorbed negative (positive) surface charges increases beyond a certain number for an n-type (p-type) semiconductor, the Fermi level crosses the middle of the bandgap and the surface region becomes p-type (n-type). This is called an inversion layer. Formation of an inversion layer is not always possible; if the dominant charge carriers in the inversion layer are annihilated faster than they are (thermally)

\(^7\) 1 monolayer (ML) corresponds to ~10\(^{15}\) atoms/cm\(^2\).
generated, no free charge carriers are present and the surface remains insulating. A deep depletion layer is then formed. Deep depletion layers are fairly common in photoelectrode materials with a relatively large bandgap (>2 eV) because generation of minority carriers is difficult in these materials. They can also be formed in the presence of surface-adsorbed species that consume the minority carriers through fast oxidation or reduction reactions. An accumulation layer can form when an excess of positive (negative) charges is adsorbed at the surface of an n-type (p-type) semiconductor. To compensate these surface charges, free majority carriers will accumulate near the surface, forming the accumulation layer.

Whereas the charge density in a depletion layer is limited by the concentration of ionized dopant ions (usually $10^{16} - 10^{18}$ cm$^{-3}$), the concentration of electronic
charge carriers is limited by the effective density of states and can reach values up to $~10^{20}$ cm$^{-3}$. Due to this high density, the widths of the accumulation layer and the thin inverted surface region for an inversion layer are usually in the order of 1–15 nm, i.e., much thinner than that of a typical depletion layer. The width of a deep depletion layer can be evaluated by the same expression as that for a normal depletion layer. The expressions for the thickness of the inversion and accumulation layers are seldom necessary in photoelectrochemical water splitting – they can be found in the literature [8, 41]. Finally, it should be noted that all four space charge layers can, in principle, be formed by applying an external potential to a semiconductor photoelectrode. The effect of applying a potential is discussed in the next section.

### 2.6 The Semiconductor–Liquid Junction

When a metal oxide semiconductor is exposed to humid air, hydroxylation of the surface occurs as described in Sect. 2.5.1. A slightly more complicated situation arises when the semiconductor is immersed in an aqueous solution. Depending on the pH of the solution, specific adsorption of protons and/or hydroxide species affect the charge distribution at the semiconductor/electrolyte interface, and thereby also the potential distribution. Analysis of the potential distribution allows us to construct an energy diagram of the complete PEC cell. This diagram is indispensable for a thorough understanding of the thermodynamics of PEC devices, as well as the charge transfer processes that make it work.

#### 2.6.1 Surface Hydroxylation

When a semiconductor is immersed in an aqueous solution, H$^+$ and OH$^-$ ions in the solution will continuously adsorb and desorb from the surface. A dynamic equilibrium will be established, which can be described by the following protonation and deprotonation reactions:

\[
M - OH \underset{k_d}{\overset{k_a}{\rightleftharpoons}} MO^- + H_{aq}^+ \tag{2.43}
\]

\[
M - OH + H_{aq}^+ \overset{k_b}{\underset{k_d}{\rightleftharpoons}} M - OH_2^+ \tag{2.44}
\]

The equilibrium of these reactions depends on the pH of the solution and the Brønsted acidity of the surface. Depending on these conditions, the net total charge adsorbed at the surface will be positive, zero, or negative. The pH at which the net adsorbed charge is zero is called the point of zero charge (PZC) of the semiconductor.
Note that the reactions (2.43) and (2.44) are only appropriate for hydroxylated semiconductor surfaces that are amphoteric, which is indeed true for most metal oxide photoelectrodes.

2.6.2 The Semiconductor/Electrolyte Interface

The structure of the semiconductor/electrolyte interface is illustrated in Fig. 2.15. The charges at the surface consists of electrons (or holes) trapped in surface states, plus the specifically adsorbed ions described in reactions (2.43) and (2.44). The counter charges are provided by ionized donors or acceptors in the solid, plus an accumulation of oppositely charged ions in the solution. Due to the large dipole moment of water, all ions in the solution are surrounded by a solvation cloud of water molecules. This cloud prevents them from approaching the surface closer than a few Ångströms (see Fig. 2.15). The region between the specifically adsorbed ions and the closest ions in the solution is called the Helmholtz layer. The potential drop across this layer is given by

---

8 "Amphoteric" means that the semiconductor surface can either donate or accept a proton, i.e., it can act both as a Brønsted acid and as a Brønsted base.
\[ V_H = \xi d = \frac{Q_S d}{\varepsilon_0 \varepsilon_r}. \] (2.45)

The width of the Helmholtz layer is \( \sim 2-5 \ \text{Å} \), and on such a small length scale the dielectric constant of water is estimated to be \( \sim 6 \), i.e., much smaller than its bulk value of 80 [42]. For a surface charge of \( Q_S = 10^{13} \ \text{cm}^{-2} \) (Fig. 2.13), the potential drop across the Helmholtz layer is typically in the order of \( \sim 0.1-0.5 \ \text{V} \). The Helmholtz capacitance is 10–20 \( \mu \text{F/cm}^2 \) [42].

During adsorption and desorption from the semiconductor surface, the ions gain or lose energy while crossing the Helmholtz layer due to the Helmholtz voltage. This leads to the following equilibrium constants for reactions (2.43) and (2.44):

\[
\frac{[\text{M} - \text{O}^-][\text{H}_\text{aq}^+]}{[\text{M} - \text{OH}]} = k_a = \exp \left( -\frac{(\Delta G_a - eV_H)}{kT} \right) \]

\[ (2.46) \]

\[
\frac{[\text{M} - \text{OH}_2^+]}{[\text{M} - \text{OH}][\text{H}_\text{aq}^+]} = k_b = \exp \left( -\frac{(\Delta G_b + eV_H)}{kT} \right). \]

\[ (2.47) \]

Making use of the fact that \( V_H = 0 \) and \( [\text{M} - \text{O}^-] = [\text{M} - \text{OH}_2^+] \) at pH = PZC, the Gibbs free energies can be eliminated to yield the following expression for \( V_H \):

\[
V_H = \frac{2.3kT}{e} \log \left( \frac{[\text{M} - \text{O}^-]}{[\text{M} - \text{OH}_2^+]} \right)^{1/2} + \frac{2.3kT}{e} (\text{PZC} - \text{pH}). \]

\[ (2.48) \]

As seen above, the excess charge at the surface is usually very small, in the order of 1\% of a monolayer or less. Hence, \( [\text{M} - \text{O}^-] \approx [\text{M} - \text{OH}_2^+] \) and (2.48) can be simplified to

\[
V_H = \frac{2.3kT}{e} (\text{PZC} - \text{pH}). \]

\[ (2.49) \]

This important result shows that the Helmholtz potential changes with \(-59 \ \text{mV} \) \((2.3kT/e)\) per pH unit at 25°C.

### 2.6.3 The Band Diagram

Now we understand the structure of the semiconductor/electrolyte interface, we can draw a detailed band diagram for a complete PEC cell. An example is shown in Fig. 2.16 for a cell composed of a n-type photoanode and a metal counter electrode. As usual, the \( y \)-axis represent the energy of an electron at a certain point \( x \) in the cell. The energy of an electron in vacuum at infinite distance is chosen as a reference. It is important to note that the vacuum level bends in the presence of an
electric field, i.e., it follows the potential gradients that are present in the PEC cell. This may seem a bit odd for a “reference” energy, but it is a consequence of the convention that the electrochemical potential (or Fermi energy) of the electrons is drawn as a constant value when the system is in equilibrium.

We now take a closer look at the electrolyte phase in Fig. 2.16. The energy of a redox couple in an electrolyte with respect to the vacuum level ($E_{\text{vacuum}}$) is not accurately known. Theoretical and experimental estimates place the $\text{H}_2/\text{H}^+$ standard redox energy between 4.3 and 4.85 eV below $E_{\text{vacuum}}$ [43]. Usually a value of 4.5 eV is taken, but it should be realized that this value has an uncertainty of a few tenths of an eV; Bockris suggested a value of $(4.6 \pm 0.2)$ eV [42]. The value recommended by IUPAC is $(4.44 \pm 0.02)$ eV [44]. The vacuum level has been drawn as a horizontal line in the electrolyte, since no electric field is assumed to be present in this phase. This is a realistic assumption for most water splitting applications because highly concentrated electrolytes are used in order to avoid Ohmic voltage losses. For very low electrolyte concentrations (<0.1 M), there may be insufficient ions available at the outer Helmholtz plane to compensate all the adsorbed charges at the semiconductor surface. These charges are then compensated in a region that extends much beyond the outer Helmholtz plane, the so-called Gouy layer. The potential distribution in this layer resembles that of a SCR in a semiconductor and is described by the Gouy–Chapman theory. Since the Gouy layer is rarely relevant for solar water splitting applications, we refer the reader to the literature for a more detailed description [42].

Fig. 2.16  Energy diagram for a PEC cell based on a n-type semiconductor and a metal counter electrode. The vacuum energy level is taken as a reference; this is the energy of an electron in vacuum at infinite distance. The electron affinity ($\gamma$) and ionization energy (IE) are materials constants, whereas the semiconductor work function ($\phi_S$) also depends on the distance to the surface. Note that a Helmholtz layer is also present at the metal/electrolyte interface.
The key parameters for the semiconductor are the band positions, the amount of band bending, and the difference between $E_C$ and $E_F$. The latter depends on the free electron concentration in the bulk ($= \frac{N_D^+ - N_A^-}{C_0}$) and can be calculated by (2.9). The energy diagram directly shows whether a certain reduction or oxidation reaction at the semiconductor surface is thermodynamically possible. In the example of Fig. 2.16, photogenerated holes will be able to oxidize water if the $\text{H}_2\text{O}/\text{O}_2$ redox energy is positioned above the top of the valence band. Similarly, only electrons with an energy above the redox energy can reduce the corresponding species in the electrolyte. By measuring the potential (i.e., the Fermi level) of the semiconductor with respect to that of a reference electrode, its reduction or oxidation power can be determined. The advantage of using a reference electrode for this is that the measured potential difference does not depend on the amount of current that flows through the cell. In contrast, the potential difference between the semiconductor and the metal counter electrode depends on $V_H$ at the metal electrode (i.e., the overpotential), which depends on the current in an unknown way.

Because of the high concentration of free electrons in the metal counter electrode, the SCR inside the metal is extremely thin (~1 Å) and can therefore be ignored. The structure of the Helmholtz layer at the metal/electrolyte interface is similar to that described for the semiconductor, and the capacitance is also in the order of 10–20 μF/cm². As already hinted in the previous paragraph, the potential drop across the Helmholtz layer depends on the kinetics of electron transfer across the interface. This is because any overpotential applied to a metal electrode must fall across the Helmholtz layer. This is quite different from the case of an external potential applied to a semiconductor electrode, which is discussed in the next section.

### 2.6.4 Applying A Bias Potential

The operation of a PEC cell can be influenced by applying an external bias potential to the semiconductor. When the bias is applied with respect to a reference electrode, the potential difference will be distributed over the space charge layer and the Helmholtz layer. These layers act as two capacitances in series [41]:

$$\frac{1}{C_{\text{tot}}} = \frac{1}{C_{SC}} + \frac{1}{C_H}.$$  \hspace{1cm} (2.50)

Both layers have the same charge $Q$ associated with them (this is the charge that is located at the inner Helmholtz plane), and since $C = Q/V$ the potential distribution is given by

$$\frac{\Delta V_{SC}}{\Delta V_H} = \frac{C_H}{C_{SC}}.$$  \hspace{1cm} (2.51)
Because \( C_H \gg C_{SC} \) (see above), any change in applied bias will fall across the depletion layer of the semiconductor.\(^9\) This important conclusion is illustrated in Fig. 2.17. As shown in this figure, applying a positive bias to an n-type semiconductor results in an increase of the depletion layer. For a p-type semiconductor, a negative bias is required to increase the depletion layer. The ability to change the band bending in a semiconductor electrode is immensely useful in the characterization of these materials, as we see in later chapters.

### 2.6.5 pH Dependence of The Band Edges

While \( V_H \) is not affected by the applied bias potential, it does depend on the pH of the solution as described by (2.49). This means that the positions of the band edges shift with \(-59 \text{ mV per pH unit}\) with respect to the redox potentials in the electrolyte. At first sight, this seems to be a very useful property: if reduction of a certain species is not possible because the conduction band is too low in energy, one might consider increasing the pH toward more alkaline values to make \( V_H \) more negative and raise \( E_C \) with respect to \( E_{\text{redox}} \) (a consequence is of course that this also increases \( E_V \), so it reduces the oxidation power of the semiconductor). Unfortunately, this strategy does not help for water splitting because the reduction and oxidation potentials of water also depend on the pH via the Nernst equation. For reactions (2.1) and (2.2) in Sect. 2.2, the Nernst equations read:

---

Note that in highly doped semiconductors (>10\(^19\) cm\(^{-3}\)) and metals, \( C_{SC} \) can exceed \( C_H \) so that any change in the applied potential will fall across the Helmholtz layer instead of the depletion layer.
Hence, the reduction and oxidation potentials depend in the same way on the pH as the band positions, i.e., by $-59$ mV per pH unit. This means that the band positions of most metal oxides are fixed with respect to the water redox potentials.

Knowledge of the band edge positions at the surface is very useful, since they determine the maximum reduction and oxidation potentials of the photogenerated electrons and holes in a semiconductor. Figure 2.18 shows the band edge energies for several important semiconductors, together with the redox potentials for water splitting. The diagram is valid for a pH of 14. As outlined above, the band positions of most metal oxides show the same pH dependence as the water reduction and oxidation potentials, so relative positions remain the same at other pH values. This is, however, not the case for semiconductors and redox couples that do not show a $-59$ mV/pH dependence (MoS$_2$, Ce$^{3+}$/Ce$^{4+}$, Ag$^+$/Ag, etc.).

It should be noted that the actual reduction power of the electrons in a photoanode is always less than that suggested by the conduction band edge.
The same is true for the oxidative power of holes generated in a photocathode. This is because the photogenerated carriers lose some of their energy while traversing the SCR. Some more energy is lost when they leave the majority carrier band of the semiconductor and enter the metal counter electrode at its Fermi level. For an n-type semiconductor, the total energy loss is given by \( E_{\text{loss}} = e\phi_{\text{SC}} + (E_C - E_F) \), and typically amounts to 0.1–0.5 eV.

### 2.6.6 The Flatband Potential

The position of the band edges with respect to the redox potentials in the electrolyte is conveniently expressed by the so-called “flatband potential”, \( \phi_{\text{FB}} \). As the word suggests, this is the potential that needs to be applied to the semiconductor to reduce the band bending to zero. The flatband situation is illustrated in the right-hand panel of Fig. 2.17. It is important to realize that the flatband potential denotes the position of the Fermi level of the semiconductor with respect to the potential of the reference electrode. This means that \( \phi_{\text{FB}} \) is slightly below the conduction band edges shown in Fig. 2.18, and that it accurately reflects the thermodynamic ability of an n-type semiconductor to reduce water to hydrogen.

Several techniques can be used to determine the flatband potential of a semiconductor. The most straightforward method is to measure the photocurrent onset potential, \( \phi_{\text{onset}} \). At potentials positive of \( \phi_{\text{FB}} \) a depletion layer forms that enables the separation of photogenerated electrons and holes, so one would expect a photocurrent. However, the actual potential that needs to be applied before a photocurrent is observed is often several tenths of a volt more positive than \( \phi_{\text{FB}} \). This can be due to recombination in the space charge layer [45], hole trapping at surface defects [46], or hole accumulation at the surface due to poor charge transfer kinetics [43]. A more reliable method for determining \( \phi_{\text{FB}} \) is electrolyte electroreflectance (EER), with which changes in the surface free electron concentration can be accurately detected [47]. The most often used method, however, is Mott–Schottky analysis. Here, the \( 1/C_{\text{SC}}^2 \) is plotted as a function of the applied potential \( \phi_A \), and the value of the flatband potential is given by the intercept on the potential axis. This can be seen from a more commonly used version of equation, (2.42), in which \( \phi_{\text{SC}} \) is replaced by \( \phi_A - \phi_{\text{FB}} \):

\[
\frac{1}{C_{\text{SC}}^2} = \frac{2}{e_0\varepsilon_\varepsilon N_D A^2} \left( \phi_A - \phi_{\text{FB}} - \frac{kT}{e} \right). \tag{2.54}
\]

Any change of the potential across the Helmholtz layer is directly reflected in the flatband potential, so it is expected to shift by \(-59 \text{ mV}\) per pH unit. This has indeed been experimentally observed for most metal oxides (e.g., TiO2, ZnO, Fe2O3, SnO2, WO3) and even some nonoxide semiconductors (Ge, GaAs, GaAs1-xP) [41, 48].
2.6.7 Surface and Interface States

Illumination of a semiconductor can lead to a change in the flatband potential, even at constant pH. This can be explained by the presence of surface states that trap photogenerated minority carriers [49]. The change in charge density at the interface (more specifically: the inner Helmholtz plane) results in a change in the Helmholtz potential and, therefore, the flatband potential. This is usually referred to as “unpinning of the band edges.” The change in flatband potential can be used to determine the density of surface states involved:

$$\Delta Q_{SS} = C_H \left( \phi_{FB}^{\text{light}} - \phi_{FB}^{\text{dark}} \right).$$

The presence of surface states can also affect the band bending in the dark. If a high density of partially filled mono-energetic surface states is present, the Fermi level will be located somewhere halfway the surface state DOS. Any change in pH or applied potential will then be accommodated by a change in the occupation of this state and a concomitant change in the potential across the Helmholtz layer. If the surface state DOS is sufficiently high, the Fermi level will show only a very small shift, and the amount of band bending in the semiconductor remains constant. This is called “Fermi level pinning”. It is fairly common in covalent semiconductors, where it can be recognized by a different amount of band bending upon forming a metal–semiconductor Schottky contact that would be expected from the work function difference between both phases. It can also be observed as a flat region in a Mott–Schottky plot because the capacitance of the depletion layer will remain constant, while the Fermi level crosses the surface state energy. The physical origin of surface or interface states can have structural defects (dangling bonds, vacancies, etc.) or chemisorbed species from an electrolyte solution. In metal oxide semiconductors, the surface state DOS is generally small and Fermi level pinning is rarely observed. However, it is important in tandem devices or photocatalysts based on a junction between covalent and metal oxide semiconductors, e.g., n-Si/n-TiO₂ [50].

The chemical nature of surface states depends strongly on the bonding character of the material. Covalent materials have dangling bonds that can act as electron donors or acceptors – these are called Shockley surface states. Moreover, overlap between neighboring bonds can result in the formation of coupled states that have bonding (donor) and antibonding (acceptor) character. In ionic materials, so-called Tamm states are formed by coordinatively unsaturated ions at the surface. For example, lattice oxygen at the surface (O₂⁻) generally forms a deep donor state just above the valence band, while lattice metal ions (M⁺) can form deep acceptor states just below the conduction band [41]. In addition to these intrinsic surface states, extrinsic surface states can be formed by the adsorption of ions from the solution.
In the discussion so far, electron transfer from the semiconductor to the electrolyte is assumed to be possible if the energy of the electron in the solid exceeds the reduction energy of the species in the solution (and vice versa). Unfortunately, this simple picture does not hold for large energy differences. Moreover, it does not give any information on the kinetics of electron transfer. A more detailed theory of electron transfer was developed in the late 1950s by Marcus [51, 52], who won the 1992 Nobel Prize in Chemistry for his contributions in this field. His theory is mainly based on classical physics, and is applicable to both homogeneous systems and reactions at electrodes. Quantum-chemical extensions of the theory were later made by the Russian scientists Levich, Dogonadze, and Kuznetsov [43]. The most useful contribution to electrochemistry was made in the early 1960s by Gerischer [53]. Whereas Marcus derived his theory using configurational coordinate diagrams of the Gibbs free energy of the species involved, Gerischer approached the problem using electronic energy levels. We follow his approach in the description below.

A common element in all versions of electron transfer theory is the concept of fluctuating energy levels in the electrolyte. In polar solvents, such as water, every ion in the electrolyte is surrounded by a cloud of oriented solvent molecules (cf. Fig. 2.15). In addition to these nearest-neighbor solvent molecules in the so-called inner sphere, the ion also has Coulomb interactions with polar solvent molecules and other ions that are farther away in the outer sphere. Due to the thermal motions of the solution, the energy level of the central ion (as seen from the electrode) will fluctuate. Moreover, after reduction or oxidation of the ion, the surrounding solvent molecules and ions will rearrange themselves to accommodate the different charge on the ion. The electrostatic energy required for this reorganization is given by $2\lambda$, where $\lambda$ is called the reorganization energy. It can be shown that the chance for a reduced or oxidized species to have a certain energy $E$ is given by [53]:

$$W_{\text{ox}}(E) = \frac{1}{\sqrt{4\pi\lambda kT}} \exp\left(\frac{-(E - E_{\text{redox}}^0 + \lambda)^2}{4kT\lambda}\right)$$

$$W_{\text{red}}(E) = \frac{1}{\sqrt{4\pi\lambda kT}} \exp\left(\frac{-(E - E_{\text{redox}}^0 - \lambda)^2}{4kT\lambda}\right).$$

Here, $E_{\text{redox}}^0$ is the standard redox potential for the species involved. The distribution of these fluctuating energy levels is shown in Fig. 2.19. It should be emphasized that this is a probability distribution – it should not be confused with the DOS, which is a fundamentally different concept (only one level exists at any one time for every ion, so optical transitions between the redox levels are not possible). The reorganization energy can be quite large compared to the bandgap of the semiconductor,
with values ranging from 0.3 eV to more than 1 eV. At room temperature, the width of each distribution curve (at half height) is given by $\Delta E \approx 0.53 \sqrt{\lambda}$ (in eV).

For semiconductors, charge transfer takes place via either the bottom of the conduction band or the top of the valence band. In certain cases, charge transfer via surface states can also occur. Although Fig. 2.19 is drawn for the case of electron transfer to the valence band, which will only be significant if photogenerated holes are present, similar pictures can be made for other pathways, e.g., $E_{\text{red}} \rightarrow E_{\text{C}}$, $E_{\text{C}} \rightarrow E_{\text{ox}}$, or $E_{\text{V}} \rightarrow E_{\text{ox}}$. An important conclusion from Fig. 2.19 is that the probability of electron transfer actually decreases if $E_{\text{red}}$ is too far above $E_S$.

This is markedly different from the behavior of metal electrodes, which show a continuous increase in current with applied potential.

The actual charge transfer process takes place via tunneling. Since tunneling is an iso-energetic process, it requires the fluctuating energy level of the species in the electrolyte to be, at a certain moment in time, equal to the energy of the electron or hole in the semiconductor. During electron transfer, the energy level of the ion is assumed to remain constant. This assumption (the so-called Frank–Condon principle) is justified, since electron transfer processes occur on a much faster time scale than ionic reorganization. The transfer rate is then proportional to the concentration of the relevant carrier in the semiconductor, the DOS of the electrolyte species ($D_{\text{ox,red}} = c_{\text{ox,red}}W_{\text{ox,red}}(E)$), and a quantum-mechanical tunneling coefficient. With this approach, the anodic (+) and cathodic (−) valence and conduction band currents can then be written as [53]:

$$ j^+_{\text{V}} = e k^+_p S_{\text{C red}} W_{\text{red}}(E_{\text{V}}^S) \quad j^-_{\text{V}} = e k^-_N V c_{\text{ox}} W_{\text{ox}}(E_{\text{V}}^S) \quad (2.58) $$

$$ j^+_C = e k^+_C N_{CC} W_{\text{red}}(E_{\text{C}}^S) \quad j^-_C = e k^-_C n_{SC} c_{\text{ox}} W_{\text{ox}}(E_{\text{C}}^S). \quad (2.59) $$

Fig. 2.19  Energy level model for electron transfer from a redox species in the electrolyte to a photogenerated hole in the semiconductor valence band. The density of states $D_{\text{red}}$ depend on the concentration of the reductant species according to $D_{\text{red}} = c_{\text{red}}W_{\text{red}}(E)$. A large overlap of the redox DOS with the energy of the hole (left) gives higher currents than a small overlap (right).
Here, $N_C$ and $N_V$ are the effective density of states in the semiconductor, and $n_S$ and $p_S$ are the surface concentrations of free electrons and holes in the conduction and valence bands, respectively. Further analysis leads to the following expressions for the overall conduction and valence band currents as a function of the overpotential $\eta$ in the presence of a redox species:

$$j_V = j_0^V \left( \exp\left(\frac{-e\eta}{kT}\right) - 1 \right) \quad \text{and} \quad j_C = -j_0^C \left( \exp\left(\frac{-e\eta}{kT}\right) - 1 \right).$$

(2.60)

It is instructive to compare this with the Butler–Volmer equation for the current at a metal electrode:

$$j = j_0 \left( \exp\left(\frac{(1 - \alpha)e\eta}{kT}\right) - \exp\left(-\alpha e\eta\right) \right).$$

(2.61)

Here, $\alpha$ is the so-called transfer coefficient, which is $\sim 0.5$ for metal electrodes. A Tafel plot of $\log(j/j_0)$ vs. $\eta$ gives a slope of $\sim 60$ mV per current decade for a semiconductor, whereas a slope of $\sim 120$ mV/decade is found for a metal. The difference is caused by the fact that the applied overpotential falls across the Helmholtz layer for a metal, and across the SCR for a semiconductor. An even more crucial difference is not immediately apparent from (2.60) and (2.61): the exchange current density for a semiconductor is much smaller than that for a metal. This is a direct consequence of the high density of states of a metal near its Fermi level, which can be 3–4 orders of magnitude higher than $N_C$ and $N_V$ in a semiconductor. In fact, the current flowing at a semiconductor photoelectrode in the dark is usually negligible (often $< 10$ nA/cm$^2$). This dark current is mainly carried by the majority carriers, i.e., electrons in the conduction band for an n-type semiconductor, and holes in the valence band for a p-type material. Under illumination, the current increases and is dominated by the transfer of minority carriers across the semiconductor/electrolyte interface. This is discussed in the next section.

### 2.7 The Photoelectrochemical Cell Under Operating Conditions

Most of the discussions above concern the properties of the semiconductor and semiconductor/electrolyte interface under equilibrium conditions, i.e., in the dark. We now explore how the system behaves under illumination.

#### 2.7.1 The Quasi-Fermi Level

Figure 2.20 shows the energy diagram of a PEC cell in the dark and under illumination. In this example, the $\text{H}_2\text{O}/\text{O}_2$ redox couple is assumed to be the most
active species and therefore dominates the electrochemical potential of the solution in the dark. The Fermi levels of the semiconductor and metal, which are electrically connected, adjust to a value close to $E_{ox}$. Upon illumination, electron–hole pairs are created and the Fermi level increases with $\Delta V_{\text{photo}}$, the internal photovoltage. Since the system is no longer in equilibrium, particularly in the SCR where the electrons and holes are generated, the use of a single Fermi level is no longer appropriate. Instead, the concept of quasi-Fermi levels is more useful. The quasi-Fermi levels are a direct measure of the concentration of electrons and holes at a certain point $x$ in the semiconductor, and are defined as:

$$n = n_0 + \Delta n = N_C e^{-(E_C - E_{F,n})/kT}$$

$$p = p_0 + \Delta p = N_V e^{-(E_V - E_{F,p})/kT}.$$  \hspace{1cm} (2.62) \hspace{1cm} (2.63)

Here, $n_0$ and $p_0$ are the equilibrium carrier concentrations in the dark, and $\Delta n$ and $\Delta p$ are the additional carriers created by illumination. For an n-type semiconductor, $n = n_0 + \Delta n \approx n_0$ and $p = p_0 + \Delta p \approx \Delta p$ so that $E_{F,n}^\infty$ remains horizontal whereas $E_{F,p}^\infty$ departs from the bulk Fermi level in the active region.

The quasi-Fermi level is often interpreted as a thermodynamic driving force. Whether or not this is appropriate is a matter of some debate. While it may provide useful insights, the concept has been derived from kinetic arguments, and can at best provide a quasi-thermodynamic description. Whereas true equilibrium thermodynamics are universally valid, the predictive value of the quasi-Fermi level as a thermodynamic driving force for, e.g., chemical reactions may depend on the reaction mechanism. One particular aspect to be noted in this respect is that the quasi-Fermi level definitions in (2.62) and (2.63) only consider electrons and holes in the conduction and valence bands. They do not account for any changes in the

---

**Fig. 2.20** Band diagram for a PEC cell based on an n-type semiconducting photoanode that is electrically connected to a metal counter electrode; in equilibrium in the dark (left) and under illumination (right). Illumination raises the Fermi level and decreases the band bending. Near the semiconductor/electrolyte interface, the Fermi level splits into quasi-Fermi levels for the electrons and holes.
occupation of surface or bulk defect states, which limits their predictive value for reactions in which these states are involved.

### 2.7.2 Photocurrent–Voltage Characteristics

The theoretical description of semiconductor photocurrent–voltage characteristics have received considerable attention in the literature. One of the most often used models was reported by Gärtner [45], who derived the following expression for the photocurrent in a semiconductor under reverse bias:

\[ j_G = j_0 + e\Phi \left( 1 - \frac{\exp(-\alpha W)}{1 + \alpha L_p} \right). \]  

(2.64)

Here, \( \Phi \) is the incident light flux, \( \alpha \) is the absorption coefficient (assuming monochromatic illumination), \( W \) is the depletion layer width, \( L_p \) is the hole diffusion length, and \( j_0 \) is the saturation current density. The model assumes that there is no recombination in the SCR and at the interface.

Gärtner’s model was improved by Reichman [45], who used more appropriate boundary conditions for his derivation of the total valence band photocurrent in an n-type semiconductor:

\[ j_V = j_G - j_0 \exp\left( -\frac{e\eta}{kT} \right) \frac{1}{1 + \frac{e\eta}{kT} \exp\left( -\frac{e\eta}{kT} \right)}. \]  

(2.65)

Here, \( j_0^V \) is the hole transfer rate at the interface, and \( j_0 \) is the saturation current density, i.e., the hole current in the valence band at \( x = W \) when \( \Phi = 0 \). \( \eta \) is the overpotential, which is defined as the difference between the applied potential and the open-circuit potential under illumination. Reichman’s model includes the possibility of recombination in the SCR, which he showed could become important when the photovoltage becomes large enough for the bands to flatten. Another useful aspect of (2.65) is that it can be used to model the effect of the slow hole transfer kinetics that are often observed for metal oxides such as \( \alpha \)-Fe\(_2\)O\(_3\) [54].

The value of the saturation current density, \( j_0 = (eL_pN_CN_V/\tau_D)\exp (-E_g/kT) \), plays an important role in Reichman’s model, especially for narrow-bandgap semiconductors with large hole diffusion lengths and low donor densities. If the bandgap exceeds \(~1.8\) eV, which is the case for almost all metal oxides, \( j_0 \) becomes negligible and (2.64) and (2.65) give almost indistinguishable results.

At about the same time as Reichman, Jarrett also published a model that explicitly included recombination in the SCR [55]. For a given impurity concentration, the numerical solutions for the photoconversion efficiency he obtained are completely defined by three parameters: the optical absorption coefficient, the minority carrier lifetime in the depletion layer, \( \tau_R \), and the charge transfer rate at
the surface. In Jarrett’s model the amount of recombination in the space charge is expressed as \((t_\text{t}/t_\text{R})\), where \(t_\text{t}\) is the time needed for the minority carriers to traverse the SCR. The usefulness of this approach is demonstrated in Chap. 9, where Jarrett’s model is used to establish design criteria for photoelectrodes based on low-mobility (i.e., small polaron) semiconducting metal oxides.

It should be emphasized that the previous models only describe the current due to minority carriers. Under near-flatband conditions, the majority carriers start to contribute to the overall current. This can be recognized by an increase in the dark current, and under those conditions (2.64) and (2.65) are no longer useful. Another implicit assumption is the absence of mass transport limitations in the electrolyte, but this is rarely an issue in photoelectrochemistry due to the high electrolyte concentrations and modest efficiencies reported thus far. Finally, it should be realized that the models described above cannot account for recombination at the semiconductor/electrolyte interface. Wilson published an extension of Gärtner’s original model that included interfacial recombination effects [56]. Although this somewhat complicated model is not often used because it does not account for the dark current and recombination in the space charge, it is of value for metal oxides that show extensive surface recombination.

2.7.3 Reaction Mechanisms

One of the main bottlenecks for water splitting is the oxidation half-reaction, which is a complicated process involving four separate electron transfer steps. Little is known about the exact molecular reaction mechanism of water oxidation on metal oxide surfaces. The only exception is TiO2, which has been extensively investigated. Early work suggested that surface Ti–OH\(^{-}\) groups are photo-oxidized to Ti-OH\(^{+}\) radicals, which further react to adsorbed H\(_2\)O\(_2\) species and finally to O\(_2\) [57]. Nakato et al. proposed an alternative mechanism based on a nucleophilic attack of a H\(_2\)O molecule on a surface-trapped hole on a lattice oxygen site. The proposed reaction mechanism is shown in Fig. 2.21 [58, 59]. This mechanism served as a basis for recent theoretical DFT studies on the water oxidation reaction at various metal oxide surfaces [60–62]. For TiO\(_2\), the reaction of water molecules with surface vacancies to form adsorbed hydroxyl groups was proposed to be the rate-determining step [61]. In contrast, the rate-limiting step for WO\(_3\) appears to be the transfer of a proton from the surface adsorbed hydroxyl group to the electrolyte [62]. Although there remains some uncertainty about the actual molecular reaction mechanisms, these theoretical studies can provide valuable pointers for rational photoanode design [60].

Since the overpotential for water oxidation can easily exceed 0.6 V for a metal oxide photoanode, co-catalysts are often used to enhance the reaction kinetics. IrO\(_2\) and RuO\(_2\) are well-known oxidation catalysts, but they are either prohibitively expensive or prone to photocorrosion [63]. Recently, low-cost and stable
alternatives based on, e.g., cobalt have been found [64, 65]. Such catalysts have indeed been reported to enhance the efficiency of hematite photoanodes [66].

The reduction of water is a much simpler two-electron process. Many n-type semiconductors can evolve hydrogen at small overpotentials and without a co-catalyst. In contrast, p-type photoelectrodes generally require a large overpotential for the onset of cathodic photocurrents, and noble metal co-catalysts are often required [43]. Few p-type semiconductors have a conduction band that is well above the hydrogen reduction energy. Notable exceptions are SiC [67] and Cu₂O [68]. An important advantage of p-type photoelectrodes is that anodic decomposition, which is often observed for n-type photoelectrodes, does not occur. Having said this, it should be noted that the stability of many p-type oxides (Cu₂O, delafossites) in aqueous environments is notoriously poor.

2.7.4 Photocorrosion

Photocorrosion occurs when the photogenerated holes do not oxidize water, but the semiconductor itself. The general form of this anodic decomposition reaction can be written as

$$
\text{M}_x\text{O}_y + \frac{2y}{x} \text{h}^+ \overset{\text{electrolyte}}{\rightleftharpoons} \text{xM}^{(2y/x)+} + \frac{y}{2} \text{O}_2. \quad (2.66)
$$

A well-known oxide that suffers from severe photocorrosion in aqueous environments is ZnO. For p-type oxides, cathodic photoreduction can lead to the formation of metal deposits:

$$
\text{M}_x\text{O}_y + 2\text{e}^- + 2\text{yH}^+_{(aq)} \overset{\text{electrolyte}}{\rightleftharpoons} \text{xM}^{(s)} + \text{yH}_2\text{O}. \quad (2.67)
$$
In an alkaline environment, the reaction can be written as

\[ M_xO_y + 2ye^- + yH_2O \xrightleftharpoons{\text{electrolyte}} \ xM(s) + 2yOH_{(aq)}, \]  

(2.68)

This has been observed for, e.g., Cu_2O. Alternatively, protons can enter the metal oxide lattice when free electrons are available at or near the surface, a process known as intercalation. It is worth mentioning that III–V semiconductors, such as GaAs and GaP, are prone to anodic photocorrosion, but appear to be relatively stable against cathodic photocorrosion in practice [69]. Whether or not these highly undesired reactions occur depends on their equilibrium potentials, which are in turn related to the Gibbs free energy of the reactions, cf. (2.5). If the energy level for anodic decomposition, \( E_{p,\text{decomp}} \), is below the valence band maximum, no anodic decomposition can occur. Conversely, if \( E_{n,\text{decomp}} \) is above the conduction band minimum, no cathodic photocorrosion can occur [41]. An overview of the anodic and cathodic decomposition potentials for various semiconductors is shown in Fig. 2.22. The actual values of the decomposition potentials depends on the concentration of the reaction products via the Nernst equation, so that they may depend on the pH of the solution [71]. A somewhat dated but useful overview of the stability of various metal oxides against photoanodic decomposition as a function of pH has been given by Scaife [72]. While some metal oxides are thermodynamically stable against cathodic photocorrosion (e.g., TiO_2), most – if not all – known metal oxides are thermodynamically unstable against anodic photocorrosion. This is also evident from the positions of \( E_{p,\text{decomp}} \) in Fig. 2.22. This is even the case for n-type photoanodes that are found to be very stable, such as TiO_2 and \( \alpha \)-Fe_2O_3; decomposition of these materials is simply inhibited by slow reaction kinetics. If a material shows signs of photocorrosion,
the addition of suitable co-catalysts may alleviate the problem to some extent by favoring the water oxidation route. In practice, photocorrosion is difficult to predict and one usually has to put new materials to the test under actual operating conditions.

2.8 Energy and Quantum Conversion Efficiencies

The energy conversion efficiency is one of the key performance indicators of the PEC cell. Assuming that all the photogenerated electrons and holes are used for the water splitting reaction, the overall solar-to-hydrogen efficiency is given by

$$\eta_{\text{STH}} = \frac{P_{\text{out}}}{P_{\text{light}}} = \frac{j_{\text{photo}} (V_{\text{redox}} - V_{\text{bias}})}{P_{\text{light}}}.$$  (2.69)

$V_{\text{redox}}$ is usually taken to be 1.23 V (at room temperature), based on a Gibbs free energy change for water splitting of 237 kJ/mol. Alternatively, the enthalpy change (286 kJ/mol) is sometimes used, which corresponds to a redox potential of 1.48 V. The latter value is appropriate when the hydrogen produced will be used in a combustion process, whereas the former value is more appropriate when the hydrogen is converted into electricity in a fuel cell [69]. When the end-use of the produced H$_2$ is not clear, it is probably best to use the Gibbs free energy value since it yields more conservative efficiency numbers. Assuming AM1.5G incident sunlight ($P_{\text{light}} = 1,000 \text{ W/m}^2$) and no applied bias, the efficiency can be directly related to the photocurrent by $\eta_{\text{STH}} = 1.23 \times j_{\text{photo}}$, with the efficiency in % and the photocurrent in mA/cm$^2$. Hence, a photocurrent of ~8 mA/cm$^2$ is required to reach the 10% efficiency target for economically viable PEC cells (see Chap. 8).

It should be emphasized that $V_{\text{bias}}$ in (2.69) refers to the actual potential difference between the working and counter electrodes. Attempts to indirectly determine this potential difference from the results of a three-electrode measurement (in which the working electrode potential is reported relative to that of a reference electrode) generally leads to errors, and should be avoided (see also Chap. 3).

A more direct method to determine the solar-to-hydrogen efficiency is to measure the amount of evolved hydrogen by mass spectrometry, gas chromatography [73, 74], or by following the water displacement in an inverted burette. The STH efficiency for this approach is given by

$$\eta_{\text{STH}} = \frac{\Phi_{\text{H}_2} G_{f,\text{H}_2}^{0}}{P_{\text{light}}},$$  (2.70)

where $\Phi_{\text{H}_2}$ is rate of H$_2$ evolution at the illuminated area (mol/s/m$^2$) and $G_{f,\text{H}_2}^{0}$ is the Gibbs free energy of formation of hydrogen (237 kJ/mol).
When reporting STH efficiencies, care should be taken to ensure that (a) the water splitting reaction is stoichiometric (H₂:O₂ = 2:1), (b) no other reactions due to, e.g., sacrificial donors/acceptors occur, and (c) the light source is well specified and closely matches the AM1.5G spectrum in intensity and spectral distribution [75]. For more detailed information on STH efficiency measurements, the reader is referred to some useful guidelines reported by a US Department of Energy (DOE) working group on photoelectrochemical hydrogen production [76] and to a careful study on the influence of the type of light source used [77].

When trying to identify performance-limiting factors in PEC photoelectrodes, the quantum efficiency as a function of wavelength is a particularly useful parameter. The definition of the external quantum efficiency is the fraction of the incident photons that is converted to electrons that can be measured by the outer circuit. This is the IPCE (incident photon-to-current conversion efficiency), which is given by

\[
\text{IPCE}(\lambda) = \frac{hc}{e} \left( \frac{j_{\text{photo}}(\lambda)}{\lambda P(\lambda)} \right)
\]

(2.71)

IPCE values in excess of 80% for the photo-oxidation of water have been reported for, e.g., WO₃ [78] and for TiO₂ under UV illumination [79]. Another useful parameter is the APCE, or absorbed photon-to-current conversion efficiency. In contrast to the IPCE, the APCE also corrects for reflection losses. Often referred to as the internal quantum efficiency, it is related to the IPCE via

\[
\text{APCE}(\lambda) = \frac{\text{IPCE}(\lambda)}{A(\lambda)} = \frac{\text{IPCE}(\lambda)}{1 - R - T}
\]

(2.72)

A, R, and T are the optical absorption, reflection, and transmission, respectively. The optical absorption is not to be confused with the absorbance or optical density, which is given by

\[
\text{OD} = -\log \left( \frac{T}{1 - R} \right) = \frac{\alpha L}{2.303}
\]

(2.73)

Here, \(\alpha\) is the absorption coefficient and \(L\) is the film thickness. The APCE is the more useful parameter when evaluating recombination within the semiconductor, while the IPCE is more appropriate to evaluate device-level conversion efficiencies.

Of particular interest is the use of IPCE values measured under monochromatic irradiation to predict the photocurrent under actual solar irradiation:

\[
J_{\text{solar}} = \int (\text{IPCE}(\lambda) \times \Phi(\lambda) \times e) \, d\lambda.
\]

(2.74)

Here, \(J_{\text{solar}}\) is the total solar photocurrent in A/m² and \(\Phi(\lambda)\) is the photon flux of sunlight in photons/m²/s. The photon flux can be calculated from tabulated solar
irradiance data, $E(\lambda)$, via $\Phi(\lambda) = E(\lambda)/(hc/\lambda)$ [75]. An important underlying assumption for (2.74) is that there is a linear relationship between the monochromatic photocurrent and the light intensity. This is often the case, but there are exceptions. A super-linear increase of the photocurrent with light intensity is sometimes observed for nanostructured materials with high concentrations of relatively shallow surface or interface traps; these traps first need to be filled before the charge carriers can reach the electrolyte or the back contact [80]. Conversely, slow charge transport or charge transfer across the interface may lead to an accumulation of free charge carriers at or near the interface. This increases the amount of recombination and will result in a sublinear increase of the photocurrent with light intensity.

### 2.9 Photoelectrode Requirements and Cell Configurations

The most critical aspect of the design of a photoelectrochemical device for water splitting is the choice of suitable photoanode and/or photocathode materials. Several of the requirements imposed on these materials appear to be in conflict, and certain trade-offs have to be made. In some cases, these trade-offs can be avoided by adopting smart architectures and materials combinations. The remainder of this chapter gives a few general considerations and approaches; more detailed accounts of strategies to cope with these trade-offs are given in later chapters.

#### 2.9.1 Requirements and Trade-offs

Most of the requirements for suitable water-splitting photoanode and/or photocathode materials have already been alluded to in the previous sections of this chapter. They can be summarized as follows [81]:

- Good (visible) light absorption
- High chemical stability in the dark and under illumination
- Band edge positions that straddle the water reduction and oxidation potentials
- Efficient charge transport in the semiconductor
- Low overpotentials for reduction/oxidation of water
- Low cost

The spectral region in which the semiconductor absorbs light is determined by the bandgap of the material. The minimum bandgap is determined by the energy required to split water (1.23 eV) plus the thermodynamic losses (0.3–0.4 eV [82]) and the overpotentials that are required at various points in the system to ensure sufficiently fast reaction kinetics (0.4–0.6 eV) [77, 83]. As a result, the bandgap should be at least 1.9 eV, which corresponds to an absorption onset at 650 nm. Below 400 nm the intensity of sunlight drops rapidly, imposing an upper limit of
3.1 eV on the bandgap. Hence, the optimum value of the bandgap should be somewhere between 1.9 and 3.1 eV, which is within the visible range of the solar spectrum. In a thorough analysis, Murphy et al. suggested an optimum bandgap of 2.03 eV, which would lead to a solar-to-hydrogen efficiency of 16.8% [77].

The chemical stability requirement is a severe one that limits the usefulness of many photo-active materials. Most nonoxide semiconductors either dissolve or form a thin oxide layer that prevents charge transfer across the semiconductor/electrolyte interface. Oxide semiconductors are more stable, but may be prone to anodic or cathodic decomposition, as outlined in Sect. 2.7.4. The general trend is that the stability against (photo)corrosion increases with increasing bandgap. Although this conflicts with the requirement of visible light absorption, a small bandgap and good chemical stability are not necessarily mutually exclusive (although one could argue that a small bandgap is often accompanied by a high valence band energy, which would indeed result in a less stable material).

Few semiconductors fulfill the third requirement, which states that the band edges should straddle the water reduction and oxidation potentials. As shown in Fig. 2.18, the ones that do either have a bandgap that is too large (e.g., SrTiO$_3$, SiC) or are unstable in aqueous solutions (Cu$_2$O, CdS). This figure also suggests that the band edges of nonoxide semiconductors tend to be better suited toward the reduction of water (high $E_C$), whereas those of oxide semiconductors favor water oxidation (low $E_V$). It should be mentioned that the third requirement is actually a bit more stringent than Fig. 2.18 suggests. As pointed out originally by Gerischer [84], and reiterated later by Weber [85], it is the quasi-Fermi levels that should straddle the water redox potentials. Although the interpretation of the quasi-Fermi level as a thermodynamic driving force may not be entirely appropriate under all conditions (cf. Sect. 2.7.1), using it within the context of the third requirement seems reasonable.

The fourth requirement, that of efficient charge transport, is easily fulfilled in some materials (TiO$_2$, WO$_3$), while in others it is one of the main causes of poor overall conversion efficiencies. A particularly important example of the latter is that of $\alpha$-Fe$_2$O$_3$, which will be discussed in detail in Chap. 4. One can distinguish intrinsic and extrinsic charge transport factors. As outlined in Sect. 2.2, the electronic band structure of the material gives important clues with regard to the intrinsic charge carrier mobilities. Extensive overlap of metal 3d orbitals usually leads to high electron mobilities, whereas the overlap of O-2p orbitals determines the hole mobility in most metal oxides. Extrinsic factors, in particular shallow donors/acceptors and recombination centers, are perhaps even more important. For example, electron transport was found to be rate limiting in spray-deposited, undoped films of $\alpha$-Fe$_2$O$_3$, whereas hole transport was found to be rate limiting after donor doping the material with silicon [86]. The susceptibility to recombination is often expressed as the carrier lifetime, $\tau_R$, or the minority carrier diffusion length, $L_D$. They are related through

$$L_D \approx \sqrt{D\tau_R}.$$  (2.75)
The diffusivity of the free carriers, $D$, is related to their mobility, $\mu$, via the Nernst–Einstein equation:

$$D = \frac{kT\mu}{e}. \quad (2.76)$$

Values of $L_D$ range from several nanometers in some oxides to several hundred micrometers in high-purity silicon.

For n-type semiconductors, the fifth requirement implies that hole transfer across the semiconductor/electrolyte interface should be sufficiently fast in order to compete with the anodic decomposition reaction. More generally, interfacial charge transfer should be fast enough to avoid the accumulation carriers at the surface, as this would lead to a decrease of the electric field and a concomitant increase in electron–hole recombination. To improve the kinetics of charge transfer, catalytically active surface species can be added. Examples of effective oxygen evolution catalysts are RuO$_2$ [87], IrO$_x$ [88], and Co-based compounds [64], whereas Pt, Rh [89], Cr–Rh, RuO$_2$, or NiO$_x$ [73, 90] are usually employed as a catalysts for hydrogen evolution.

At this time, no single material has been found that meets all these requirements. Four main trade-offs can be identified:

- Bandgap ↔ stability
- Photon collection ↔ charge transport
- Recombination ↔ catalysis
- Performance ↔ cost

The bandgap vs. stability issue has been discussed above. The second trade-off is especially important for semiconductors with a small absorption coefficient, such as metal oxides with an indirect bandgap. In these materials most of the electron–hole pairs are generated far away from the surface, and the carriers recombine before reaching the interface (cf. Fig. 2.23). The third trade-off reflects the tendency of catalytically active surface sites to also act as efficient recombination centers for electronic charge carriers. Moreover, certain co-catalysts for hydrogen or oxygen evolution may also enhance the back-reaction between H$_2$ and O$_2$. A well-known example is Pt. This may be a problem for cell designs in which the cross-over of gases generated at the (photo)anode and (photo)cathode cannot be avoided. Interestingly, NiO is an efficient hydrogen evolution catalyst at which the H$_2$ + O$_2$ back-reaction does not seem to occur [73]. The final trade-off reflects the main reason for the interest in low-cost metal oxides, since the economic viability of state-of-the-art multijunction devices based on, e.g., high quality III–V materials grown by MOCVD or MBE is doubtful [91].
2.9.2 Nanostructured Electrodes

Nanostructured electrode morphologies can be used to address some of the intrinsic material’s limitations and trade-offs mentioned above. The most obvious advantage of a nanostructured morphology is the increase in specific surface area. The concomitant increase in the number of surface sites greatly enhances the overall charge transfer kinetics at the semiconductor/electrolyte interface.

A second advantage is the shorter diffusion path lengths for the photogenerated charge carriers. This is illustrated by the nanowire array photoelectrode shown in Fig. 2.24 (left), where the minority carriers only have to travel half the diameter of the nanowire in order to reach the semiconductor/electrolyte interface [92]. If transport of majority carriers is also an issue, one could employ an array of conducting wires, which are coated with a thin film of the photoactive semiconductor (Fig. 2.24, right). Here, both minority and majority carriers benefit from the short distance ($\leq d$) that they have to travel before reaching the electrolyte and conducting core, respectively. This is sometimes referred to as a “guest–host” structure, in which the light absorbing nanosized guest material is attached to a nanostructured host scaffold [93]. The advantages are not limited to the highly regular nanowire array architecture shown in Fig. 2.24; nanostructures based on dendritic or cauliflower-type morphologies or randomly packed spheres offer similar advantages [78, 93, 94]. Such structures are discussed in more detail in Chaps. 4 and 9.

Fig. 2.23 Absorption regions in a semiconducting photoanode. The intensity of the light ($\Phi$) decreases exponentially with distance from the surface (Lambert-Beer law). All charge carriers generated within a distance $L_D + W$ from the surface contribute to the photocurrent (regions I and II), whereas all carriers generated in region III recombine.

Region I:
- Absorption
- Charge separation
- Field-assisted transport (drift)

Region II:
- Absorption
- Transport by diffusion
- Holes are able to reach region I before recombining

Region III:
- Absorption + recombination

Optimal film thickness: $d = \alpha^{-1} \approx W + L_D$
A third possible advantage of using materials with nanoscale dimensions is the occurrence of quantum size effects. As first described by Brus in the early 1980s, spatial confinement of charge carriers to a volume that is less than their De Broglie wavelength results in a widening of the bandgap [9]. An approximate expression for the bandgap enlargement as a function of particle radius ($R$) is given by [95]:

$$E_g' = E_g^0 + \frac{\hbar^2 \pi^2}{2R^2} \left( \frac{1}{m_e^*} + \frac{1}{m_h^*} \right) - \frac{1.8e^2}{\varepsilon_0 \varepsilon_r R}.$$  

(2.77)

Here, $E_g^0$ is the bulk bandgap and $m_e^*$ and $m_h^*$ are the effective masses for the electrons and holes, respectively. The first term describes the quantum confinement effect, while the second (smaller) term describes the attractive Coulomb interaction between the electrons and holes. The negative sign of the second term shows that it in fact acts against the bandgap enlargement. While the resulting blue shift of the absorption spectrum is usually not desirable for PEC materials, the widening of the gap shifts the conduction band edge positively with respect to the hydrogen reduction energy. This would decrease the required bias potential and increase the solar-to-hydrogen conversion efficiency (2.69) of materials such as $\alpha$-Fe$_2$O$_3$ and WO$_3$, for which $E_C$ is located below $E^0(H_2/H^+)$ – see Fig. 2.18. Although one recent report claims a 0.3–0.6 eV upward shift of the conduction band of $\alpha$-Fe$_2$O$_3$ nanowires [96], no direct photoelectrochemical evidence for such a shift in $\alpha$-Fe$_2$O$_3$ has been reported so far.

It should be noted that quantum size effects are difficult to induce in most metal oxides. This is due to their relatively large electron and hole effective masses (cf. Table 2.1), which would require exceedingly small particle sizes (<2–3 nm) for any confinement effect to occur. ZnO is a well-known exception that shows appreciable bandgap enlargement at particle sizes below 6 nm [11, 97].
enlargement is generally asymmetric, with $\Delta E_C = \frac{\hbar^2 \pi^2}{(2R^2 m_e^*)}$ and $\Delta E_V = \frac{\hbar^2 \pi^2}{(2R^2 m_h^*)}$. This asymmetry is less evident in ZnO ($m_e^* = 0.24m_e$, $m_h^* = 0.24m_e$) than in anatase TiO$_2$ ($m_h^* = 0.8m_e$, $m_e^* = 10m_e$), for which any bandgap enlargement appears to be solely due to a shift in $E_V$ [11]. Clearly, if quantum confinement is to be used to shift the conduction band edge to higher energies without increasing the bandgap too much, one would prefer to have a semiconductor with a small electron effective mass and a large hole effective mass.

An important aspect of nanosized semiconductors is the absence of band bending. Since the number of ionized donors or acceptors in nanosized particles is very small, such particles cannot sustain a large built-in electric field (cf. (2.41) and Fig. 2.13, left). The maximum attainable potential drop within a spherical semiconductor particle with radius $R$ and a donor (or acceptor) density $N_D$ is given by [98]:

$$\phi_{SC} = \frac{-eN_D R^2}{6\varepsilon_0 \varepsilon_r}.$$  

From this expression, the band bending in a 10 nm $\alpha$-Fe$_2$O$_3$ particle ($R = 5$ nm, $\varepsilon_r = 25$) with a relatively high donor density of $10^{18}$ cm$^{-3}$ is limited to $\sim 3$ mV. This clearly shows that the electric field in the space charge of a nanoparticle cannot play a significant role in the separation of photogenerated electrons and holes. Fortunately, most carriers are nevertheless able to reach the interface since the minority carrier diffusion length is – with few exceptions – much larger than the radius of the nanoparticle. As a result, the charge transfer kinetics at the semiconductor/electrolyte interface usually dominate the overall charge separation and recombination kinetics in nanoparticles.

### 2.9.3 PEC Device Configurations

When the conduction and valence bands do not straddle the water reduction and oxidation potentials, or if the available overpotentials for water splitting are insufficient, a combination of two or more semiconductors may offer a solution. In such a device, one semiconductor acts as a regular photoanode or photocathode. The other semiconductor(s) function(s) as either a complementary photoelectrode, or as an integrated photovoltaic device that provides a bias voltage. Figure 2.25 shows some examples of possible configurations.

Configuration (A) is the one we have used throughout this chapter to illustrate the main principles of photoelectrochemical water splitting. Strontium titanate, SrTiO$_3$, is one of very few materials that have been observed to split water in this configuration [99]. However, the efficiency is less than 1% due to its large bandgap (3.2 eV). Cuprous oxide, Cu$_2$O, and p-type SiC are candidates for configuration (B). While Cu$_2$O has been found to suffer from photoreduction to metallic Cu [100, 101],
other reports indicate that it may be stable under the right conditions [68, 102]. The 4H polytype of silicon carbide is one of the few chemically stable photocathode materials able to split water, but its efficiency is also low ($E_g = 3.26$ eV) [67].

In configurations (C) and (D) a bias voltage is applied to the photoanode and photocathode, respectively, through an integrated p–n junction. Turner et al. have

![Fig. 2.25 Examples of possible PEC configurations under illumination. Top row: Standard single-semiconductor devices based on a photoanode (a) or photocathode (b) with a metal counter electrode. Middle row: Monolithic devices based on a photoanode (c) or photocathode (d) biased with an integrated p–n junction. Bottom row: p–n junction photoelectrochemical device (e), and an n–n heterojunction PEC device based on a photoanode deposited on top of a second n-type semiconductor that “boosts” the energy of the electrons (f).]
demonstrated such a monolithic tandem device based on a photocathode of p-type GaInP$_2$ in combination with a GaAs p–n junction [103]. They achieved an impressive solar-to-hydrogen efficiency of 12.4%, but the rapid degradation and high cost of the device limit its practical use. Grätzel and Augustynski used a dye-sensitized solar cell (DSSC) to bias a WO$_3$ or Fe$_2$O$_3$ photoanode [91, 104]. The photoanode absorbs the blue part of the solar spectrum ($E_g = 2.6$ eV for WO$_3$), whereas the DSSC absorbs red light. This approach is somewhat similar to configuration (C), except for the fact that the p–n junction is replaced by a separate DSSC device. The advantage is that the entire system can be made from low-cost and stable metal oxides. A remaining challenge is to address the significant light scattering of the individual (nanostructured) components, which severely limits the efficiency of the total system [105].

In configurations (E) and (F) an additional bias potential is generated by a single additional absorber instead of a p–n junction. An example of configuration (E) is a device based on n-type TiO$_2$ ($E_g = 3.2$ eV) and p-type GaP ($E_g = 2.25$ eV) [106]. The difference in bandgaps ensures that a large part of the solar spectrum is used. This is not the case for p–n homojunctions, and devices such as n-Fe$_2$O$_3$/p-Fe$_2$O$_3$ [107] will not be very efficient because of this. The n–n heterojunction of configuration (F) was first studied many years ago, and was not considered to be very useful since the photogenerated holes in the buried n-type semiconductor could not reach the photoanode/electrolyte interface [108]. More recently, however, Yang et al. showed that n-type Si underneath n-type TiO$_2$ “boosts” the energy of TiO$_2$ conduction band electrons [50]. This increases the available overpotential for the reduction of water and improves the kinetics.

Acknowledgments The author gratefully acknowledges Fatwa F. Abdi for critical reading of the manuscript, and the NWO-ACTS Sustainable Hydrogen program (project 053.61.009) and the European Commission’s Framework 7 program (NanoPEC, Project 227179) for support.

References

Chapter 3
Photoelectrochemical Measurements

Roel van de Krol

3.1 Introduction

Despite the many advances in computational materials science in the last decade, photoelectrochemical water splitting is still very much an experimental field. This is because many of the relevant properties of photoelectrodes are determined by factors that are either difficult to control or not yet well understood, which makes it difficult to predict a priori their influence on the actual performance. An example of this is the presence of certain defects, such as dislocations or impurities. These are often impossible to avoid, especially when low-temperature and low-cost synthesis methods are used. An important stage in the development process of a photoelectrode is therefore to systematically optimize the synthesis procedure in order to achieve a maximum photoresponse.

A common metric to determine the optimal photoresponse is the photocurrent density under 1 sun illumination (1,000 W/m²). To be able to compare the photocurrent for different samples, two criteria should be fulfilled. The main one is that the measured data reflect the properties of the sample, as opposed to artifacts or limitations caused by the measurement setup. For example, mass transport limitations should be avoided by careful design of the electrochemical cell, and Ohmic losses and pH fluctuations can be avoided by choosing a suitable electrolyte. For a meaningful comparison of performance numbers reported by different laboratories one should also fulfill the second criterion, which is that the measurements must be

1 Examples of optimization parameters for photoelectrode synthesis are deposition temperature, substrate cleaning method, additives used to optimize the boiling point, surface tension, water content or viscosity of a solution, postdeposition heat treatments to improve the crystallinity, dipping as-prepared samples in a precursor solution to improve interparticle contacts, etc.

R. van de Krol (✉)
Faculty of Applied Sciences, Department Chemical Engineering/Materials for Energy Conversion and Storage, Delft University of Technology, P.O. Box 5045, 2600 GA Delft, The Netherlands
e-mail: r.vandekrol@tudelft.nl

performed under the exact same conditions. While seemingly trivial, there are many performance claims in the literature that fail to meet this requirement. This can be due to lack of, e.g., proper equipment (calibrated light sources in particular), but also to missing information in the paper to which the authors compare their results. To avoid this, one should be careful to mention at least the sample area and thickness, the intensity and type of light source used, the electrolyte composition and pH, the cell geometry and counter electrode area, and the applied potential.

In addition to performance evaluations, many photoelectrochemical experiments are aimed to identify performance-limiting steps or to determine certain materials properties. Examples of the latter are donor or acceptor densities and the flatband potential of a material, which can be determined by electrochemical impedance measurements. The challenge with these measurements is that they always yield data, but that it can be difficult – and sometimes even impossible – to translate the measured data to the desired materials parameters. Carefully performed control experiments and a good basic understanding of the measurement equipment – in particular, the potentiostat and the frequency response analyzer (FRA) – are essential for obtaining meaningful results.

Much information on photoelectrochemical measurements and measurement equipment is available, but it is often scattered throughout the various papers, books, and technical notes of manufacturers. Moreover, many practical methods and tricks are passed on by communication between colleagues or between supervisors and students, and cannot be found in the literature. This chapter attempts to bring some of the basic information on photoelectrochemical measurement techniques and equipment together and aims to serve as a quick start for students and other newcomers in this exciting field.

### 3.2 The Photoelectrochemical Cell

One of the key components of a photoelectrochemical setup is the photoelectrochemical cell (further abbreviated as PEC cell) in which the photoanode or photocathode is mounted. At bare minimum, a PEC cell consists of a reservoir to hold the electrolyte, the sample to be studied (the working electrode), a counter electrode to supply current, an optically transparent window that allows the sample to be illuminated, and facilities to electrically connect both electrodes to the outside world. Other components that are often present are a reference electrode, a quasi-reference electrode (see Sects. 3.2.4 and 3.6.6), and one or more inputs and outputs for a gas circulation system and/or gas purging of the electrolyte. In addition, one may encounter a magnetic stir bar, and a membrane that separates the anode and cathode compartments in order to avoid mixing of the evolved oxygen and hydrogen gases.

---

2 Most photoelectrochemical cells used in PEC research are designed to study the properties of a single photoelectrode. PEC cells in which more than one electrode is illuminated at once (e.g., tandem cells) are somewhat more complicated and will not be considered here.
3.2.1 Cell Designs

An example of a common PEC cell design is shown in Fig. 3.1. This design is based on that of conventional electrochemical cells used for, e.g., corrosion studies. The main aspect that distinguishes the PEC cell from a standard electrochemical cell is the presence of an optically transparent window through which the sample can be illuminated. Since normal glass shows a transmission cut-off below ~350 nm, fused silica (amorphous SiO$_2$) is most often used as a window material in PEC cells. UV-grade fused silica is transparent in the entire range from UV to near-IR with a transmission $>90\%$ between 200 nm and 2.2 $\mu$m.\(^3\) It shows excellent chemical stability in both acidic and alkaline aqueous solutions (except in HF) and is relatively cheap.\(^4\) Cells such as the one shown in Fig. 3.1 are commercially available, fairly robust (despite being constructed entirely from glass), and easy to use when it comes to mounting the electrodes. There are also some disadvantages. As it is made from glass, the cell can be rather large and may require a large amount of electrolyte (>100 mL). This quickly becomes impractical when screening large numbers of photoelectrodes. Moreover, the distance between the working and counter electrodes usually exceeds 5 cm and may give rise to Ohmic voltage losses at higher current densities (see Sect. 3.2.5). A more practical issue is sample

\(^3\) Except for a small dip around ~1400 nm, depending on the purity grade.

\(^4\) A $\varnothing$50 mm, 3 mm thick window of UV-grade fused silica costs ~40 EUR (in 2010). Note that fused silica is often incorrectly referred to as “quartz,” which is the crystalline form of SiO$_2$ that shows slightly better transmission than fused silica at a much higher price.
mounting: since the sample is fully immersed in the electrolyte, it needs to be partially masked with, e.g., epoxy resin to avoid any direct contact between the electrolyte and the conducting back-contact and/or the connecting wire.

To address these issues, many groups design their own photoelectrochemical cells that are specifically optimized for the types of samples and experiments that are carried out in their laboratories. Figure 3.2 shows an example of a cell that was developed in the author’s laboratory. It has a modest volume (~30 mL), and the sample is mounted in a detachable sample insert for ease of handling. The sample insert is pressed against an O-ring seal on the cell body via ring that is threaded at the outside (R1). A similar ring (R3) is used to mount the fused silica window at the front of the cell. The sample surface is pressed against an O-ring with another threaded insert (R2). Electrical contact to the sample is made with a thin copper wire that is glued to the sample with carbon or silver paste. The wire is guided to the outside via a narrow slot (F) in the sample insert. The design ensures that the sample area wetted by the electrolyte is the same as the illuminated area. The sample can be illuminated from the front or from the back. Since the center of the sample coincides with the center of the four alignment rods, switching from front- to back-illumination can be done by rotating the cell 180° without risk of losing optical alignment. At the bottom of the cell a cylindrical hole is made to accommodate a small magnetic stirbar, and the same hole is connected to a drain at the front of the cell. The drain allows the electrolyte to be refreshed without removing the cell or the cell lid from the setup. The feedthroughs in the lid of the cell are threaded on the inside, with an O-ring seal to ensure gas tightness.

Fig. 3.2 Photoelectrochemical cell designed by the author’s group at TU Delft
A variation of this cell was designed by the LPI group at EPFL. Their “Cappuccino” cell has an even smaller electrolyte volume and is fabricated from PEEK (polyether ether ketone), a chemically resistant polymer that is easier to machine and has better mechanical properties than PTFE. Their simple but ingenious design allows two methods of sample mounting. One is by pressing the sample against an O-ring, similar to the approach shown in Fig. 3.2. In the other method, the top of the sample is clamped in a miniature steel vise. The vise is mounted at the top of the cell, so that the bottom part of the sample is fully immersed in the electrolyte. The top part of the sample should not be coated, so that the vise connects with the conducting back-contact and thereby serves as an electrical contact. This configuration is particularly useful for (meso)porous samples, which cannot be easily sealed with an O-ring. The illuminated area is defined by a circular hole at the front of the cell. The clamped sample – including the vise – can be easily lifted out of the cell and rotated 180° for back-side illumination. The Cappuccino design is particularly suited for fast screening of nanostructured samples. It is easier to handle than the design of Fig. 3.2, but with the limitations of not being gas tight and having a wetted area that is ~2–3 times larger than the illuminated area.

### 3.2.2 Sample and Substrate

From a practical point of view, the main characteristic of a photoanode or photocathode sample is whether the film is compact or porous. Compact films can be easily sealed by gently pressing them against an O-ring. The area that is exposed to the electrolyte can be accurately determined, which is essential for, e.g., donor density calculations using the Mott–Schottky relationship – see (2.54) and Sect. 3.6. Moreover, the illuminated area is the same as the wetted area, which minimizes dark current contributions to the total current. Porous films can be mounted by gluing them to a mask with a hole that defines the area to be wetted. The choice of glue or sealant is critical: it needs to be chemically resistant in the electrolyte to be used, and it also needs to have a fairly high viscosity to prevent the infiltration of the porous film in the area that is exposed to the electrolyte. In practice, this limits the choice to certain epoxy resins, or to certain thermoplastic polymers that soften between ~100 and 200°C. Examples of the latter are Surlyn® and Bynel® from Dupont, which are used to seal dye-sensitized solar cells based on mesoporous TiO₂. These sealants are sold as 25–50 μm sheets and are convenient to use, but somewhat costly. Alternatively, one can simply immerse the entire sample (except the top part) in the electrolyte, as discussed above for the Cappuccino cell. Care should be taken that the entire sample is coated in order to avoid direct exposure of the conducting substrate to the electrolyte. This is the most convenient method for photocurrent measurements on porous films. However, it is not suited for quantitative impedance analysis, since the meniscus at the sample/water/air interface results in a poorly defined sample area.
Photoelectrochemical measurements on thin film samples require a conducting substrate as a back-contact. Optically transparent substrates are usually preferred since this also allows the sample to be illuminated from the back. The substrate should form an Ohmic contact with the photoactive material to ensure uninhibited current flow. For n-type photoanodes, this requires conducting materials with a work function that is lower than that of the photoanode itself. In most cases, transparent conducting oxides (TCOs) are used. These materials have high carrier concentrations and mobilities, but still show good optical transparency (>80%) because their plasma frequencies are in the near infrared (\( \lambda_p > 1 \, \mu m \)). They are deposited as a 100–400 nm film on glass by sputtering or spray pyrolysis. Examples of TCOs are F-doped SnO\(_2\) (FTO), Sb-doped SnO\(_2\) (ATO), Sn-doped In\(_2\)O\(_3\) (ITO), and Al-doped ZnO (AZO). Table 3.1 indicates several key materials properties of important TCOs. Many other TCOs exist, but FTO and ITO are most common ones that are commercially available. Examples of suppliers are Pilkington, Asahi, Libbey–Owens–Ford, and several smaller ones such as Solaronix and Dyesol. Typical sheet resistances of commercial TCOs on glass are 8–30 \( \Omega /sq. \) ITO has a ~2 times higher conductivity than FTO, but shows poorer thermal stability and chemical resistivity [2]. Since sheet resistances are not (yet!) limiting the performance of PEC cells for water splitting, FTO is currently the conducting substrate of choice. For p-type photocathodes, high-work function back-contacts such as gold and platinum can be used. It should be noted that work function differences do not always correctly predict the formation of an Ohmic or Schottky contact [3]. For example, FTO appears to form an Ohmic contact with certain p-type materials such as Cu\(_2\)O [4].

The choice of conducting substrate becomes more difficult when postdeposition high-temperature treatments are necessary. This is often the case for complex oxides, such as perovskites or oxynitrides [9], which may require firing at temperatures above 600\(^\circ\)C in order to obtain the desired crystalline phase. This prohibits the use of float glass, which softens above ~550\(^\circ\)C. Certain types of borosilicate glass can be used up to 650\(^\circ\)C, while fused silica or sapphire can withstand continuous exposure to temperatures up to 950\(^\circ\)C. Unfortunately, the conductivity of ITO films quickly decreases above 350\(^\circ\)C. FTO and ITO/FTO coatings are stable up to 600–700\(^\circ\)C [2, 10], and may still have acceptable conductivities at higher temperatures provided

### Table 3.1  Typical materials parameters for some common TCOs [2, 5–8]

<table>
<thead>
<tr>
<th>TCO material</th>
<th>Highest conductivity (S/cm)</th>
<th>Carrier concentration (cm(^{-3}))</th>
<th>Mobility (cm(^2/V/s))</th>
<th>Work function (eV)</th>
<th>Thermal stability in air ((^\circ)C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In(_2)O(_3):Sn (ITO)</td>
<td>(~1 \times 10^4)</td>
<td>(~10^{21})</td>
<td>~40</td>
<td>4.7</td>
<td>(&lt;350)^a</td>
</tr>
<tr>
<td>SnO(_2):F (FTO)</td>
<td>(~1 \times 10^3)</td>
<td>4 \times 10^{20}</td>
<td>~30</td>
<td>4.9</td>
<td>(&lt;700)</td>
</tr>
<tr>
<td>ZnO:Al (AZO)</td>
<td>(~7 \times 10^3)</td>
<td>1.5 \times 10^{21}</td>
<td>50–100</td>
<td>4.6</td>
<td>(&lt;400)</td>
</tr>
</tbody>
</table>

\(^a\)A thermal stability of up to 900\(^\circ\)C in air has been reported for ITO on fused silica [8]
the anneal treatment is kept sufficiently short (<30 min). For higher temperatures and/or longer times one needs to switch to thin noble metal films (e.g., Pt), at the price of significantly poorer optical transmission properties.

### 3.2.3 The Counter Electrode

To avoid performance limitations, the reaction at the counter electrode should be fast and the electrode itself should have a high catalytic activity. A common materials choice is Pt, which combines good chemical stability with a very small overpotential for hydrogen evolution (~0.1 V). Ideally, the surface area should be at least two times larger than the working electrode area, and both electrodes should face each other symmetrically to avoid inhomogeneous current densities at the working electrode. For high electrolyte concentrations (typically >0.5 M, see below), the latter requirement is less stringent due to the modest voltage drop across the electrolyte solution. A common configuration is a coiled Pt wire electrode or a piece of Pt wire mesh at about ~1 cm in front of the working electrode. A circular hole in the mesh is necessary to avoid it from blocking the light. A piece of TCO glass coated with a sputtered film of Pt is a low cost and – in most cases – perfectly acceptable alternative.

### 3.2.4 The Reference Electrode

When studying the properties of the working electrode – either a photoanode or a photocathode – the applied potential is a key parameter. This potential should be measured with respect to a fixed reference potential, so that any change in the applied potential reflects a change in the working electrode alone. The counter electrode cannot be used for this, since the overpotential at the counter electrode/electrolyte interface is generally unknown\(^5\) and varies with the amount of current flowing through the cell. To avoid this current-dependence, a third electrode is added to the electrochemical cell: this is the reference electrode. The potential of the working electrode can now be measured with respect to the (fixed) potential of the reference electrode with negligible current flow through the latter.

Table 3.2 shows an overview of commonly used reference electrodes in solar water splitting research. Applied potentials are nowadays usually reported against the RHE scale for water splitting studies. Zero volt on the RHE scale reflects the

---

\(^5\)Factors that may contribute to the total overpotential include slow charge transfer across the Helmholtz layer, slow reaction kinetics due to preceding or subsequent reaction steps, mass transport limitations (diffusion, convection, migration), and removal of the solvation sheet of water molecules (dipoles) surrounding each ion.
H+/H₂ redox potential in the actual solution, irrespective of the pH. This makes it more convenient to use than the NHE scale, for which \( f(H+/H₂) = 0 \text{ V} \) only at unit activity of the H⁺ ions and the dissolved H₂ gas. Silver/silver chloride electrodes are most commonly used as reference electrodes, and have largely replaced the traditional and environmentally less friendly saturated calomel electrodes (SCEs) that are based on Hg/Hg₂Cl₂. A potential measured with respect to Ag/AgCl \( f(\text{Ag/AgCl}) \) can be converted to the RHE scale \( f(\text{RHE}) \) with the following expression:

\[
\phi_{\text{RHE}} = \phi_{\text{Ag/AgCl}} + \phi_{0}^{\text{Ag/AgCl vs. SHE}} + 0.059 \times \text{pH}.
\]  

(3.1)

Here, \( \phi_{0}^{\text{Ag/AgCl vs. SHE}} \) is the potential of the Ag/AgCl reference electrode with respect to the standard hydrogen potential (SHE). To give an example, a potential of 0.23 V vs. Ag/AgCl (satd.) in a 1 M KOH solution of pH 13.6 corresponds to a potential of 1.23 V vs. RHE.

An important selection criterion for a reference electrode is its stability in the electrolyte solution. An example of a good all-round electrode that can also be used in alkaline solutions is the XR-300 (Ag/AgCl, \( \phi_{0} = 0.198 \text{ V} \) vs. SHE) from Radiometer Analytical [11]. Other well known suppliers are Metrohm, Koslow,
several potentiostat manufacturers, and many others. It should be emphasized that all reference electrodes are delicate should be well-maintained; they should not be allowed to dry out and should be stored in the appropriate solution when not in use. One should consider having at least three identical reference electrodes in the laboratory, and check their potential differences every 2 weeks or so to confirm that the deviation between any two individual electrodes is less than ±3 mV. Larger differences may indicate that the electrode solution should be refreshed, that the frit should be declogged (by, e.g., overnight immersion in a 1 M HCl solution), or that the electrode is at the end of its life. The lifetime of a reference electrode depends on its design and manufacturer, but they typically last ~2–3 years if properly maintained. More detailed information on various types of reference electrodes can be found in Sawyer [12].

3.2.5 The Electrolyte

The electrolyte in an electrochemical cell consists of a solvent in which the active species to be reduced or oxidized are dissolved. For photoelectrochemical water splitting, the solvent and the active species are one and the same: water. However, since pure water is poorly conducting, supporting ions must be added to ensure that the desired current flow can be attained. The concentration of the supporting electrolyte should be sufficiently high to avoid large Ohmic voltage losses across the electrolyte. These Ohmic losses have to be compensated by increasing the applied bias potential, which adversely affects the overall efficiency of the photoelectrode – see (2.69). The voltage drop is given by $V_{\text{loss}} = I \times R_E$, where $I$ is the total current flowing between the working and counter electrodes, and $R_E$ is the electrolyte resistance. The latter is determined by the conductivity of the electrolyte, $\kappa$, and the cell constant, $K_{\text{cell}}$, according to:

$$R_E = \frac{K_{\text{cell}}}{\kappa}.$$  

(3.2)

For the simple geometry shown in Fig. 3.3a, the cell constant is given by the distance between the electrodes divided by the electrode area, $K_{\text{cell}} = L/A$. The cell constant depends on the geometry of the cell, and can be determined from an ac impedance measurement of $R_E$ using a standard electrolyte solution of known conductivity.

Another consequence of a non-negligible voltage drop across the electrolyte is a deviation of the measured reference electrode potential, $\phi^*_{\text{REF}}$, from the “real” reference potential, $\phi_{\text{REF}}$, as illustrated in Fig. 3.3b. To avoid this, the reference electrode needs to be placed as close as possible to the working electrode. When this is not possible due to the design of the and/cell or the size of the reference electrode, a so-called Luggin capillary can be used. By placing the opening of this capillary close to the surface of the photoelectrode (at ~0.5–1 mm), the reference electrode “samples” the electrolyte potential at this location. While Luggin capillaries are commonplace in electrochemistry, they are rarely mentioned in publications on
solar water splitting. This is because the high electrolyte concentrations used in these studies usually implies that the voltage drop across the electrolyte can be neglected. However, as we shall see below this may no longer be true when photocurrent densities reach values of a few mA/cm², and for studies on such highly efficient photoelectrodes the use of a Luggin capillary is recommended.

The electrolyte conductivity depends strongly on the type of dissolved ions and their concentrations. It is important to realize that the relationship between the conductivity and the concentration is in general non-linear. This is due to either incomplete dissociation of the anions and cations (in the case of so-called weak electrolytes) and/or ion–solvent interactions. Deviations from linearity may already occur at concentrations above 1 mmol/L, and will certainly play a role in practical water splitting applications. At very high concentrations (>1 M), ion-pair formation may result in a decrease of the conductivity with concentration. This is why an aqueous solution of KOH shows a conductivity maximum at a concentration of ~6 M. Table 3.3 gives an overview of the conductivities of several electrolyte solutions often used in PEC studies. To give a rough idea of the Ohmic losses that can be expected, Table 3.3 also includes the voltage drop across the electrolyte for a current density of 5 mA/cm². Clearly, concentrations of at least 0.5 M are required to keep Ohmic losses at an acceptable level. Moreover, potassium-based salts are preferred over sodium salts from a conductivity point of view.

In practice, aqueous H₂SO₄ or HCl solutions (0.5–1 M) are often used as an acidic electrolyte for, e.g., WO₃ [19] or TiO₂. For WO₃, NaCl has also been used in

---

6 Each ion is surrounded by a solvation sheath of water molecules (see Sect. 2.6.2). The size and charge of the central ion determine the configuration of the surrounding molecules, and this in turn determines the interaction strength of the ion with neighboring ions in the solution.
order to simulate saline conditions [20]. For photoelectrodes that require neutral or alkaline solutions, such as α-Fe₂O₃ [21], 0.5–1 M KOH or NaOH solutions are commonly used. Some metal oxides are only stable in fairly neutral environments. An example is BiVO₄, which is only stable between pH 3 and 11 and is usually studied in an aqueous solution of 0.5 M Na₂SO₄ or K₂SO₄ [22]. To avoid local pH fluctuations in this solution – which would affect the flatband potential – one should consider to add a pH buffer to the solution. An suitable choice is the well-known phosphate buffer KH₂PO₄/K₂HPO₄, which gives a pH of 6.86 at 25°C when adding 0.025 M of each species to the electrolyte [16].

To efficiently remove the evolved hydrogen and oxygen from the electrolyte solution, the electrolyte should be continuously purged with an inert gas such as nitrogen or argon. This prevents the back-reaction of dissolved H₂ and O₂ to water, and ensures that the redox potentials do not change over time (both H₂ and O₂ appear in the Nernst equations for water reduction and oxidation, see (2.52–2.53)). Continuous purging through the electrolyte solution must be avoided when the generated bubbles interfere with the light path and generate excessive noise on the photocurrent signal. If this is unavoidable, one can purge the solution thoroughly (at least 20 min) before starting the experiment, and hang the purge tube above the electrolyte to generate a blanket of inert gas above it. Stirring the solution with a magnetic stir bar greatly enhances the purging efficiency and ensures quick removal of any dissolved oxygen. Stirring to address mass transport limitations in the electrolyte is usually not necessary in view of the high ionic concentrations, and is often avoided to prevent measurement noise.

### Table 3.3 Electrical conductivities for various electrolytes and the corresponding electrolyte resistances and voltage drops for a current density of 5 mA/cm²

<table>
<thead>
<tr>
<th>Reference electrode</th>
<th>κ (Ω⁻¹ m⁻¹)</th>
<th>$R_E$ (Ω)</th>
<th>$V_{loss}$ @ 5 mA/cm² (mV)</th>
<th>T (°C)</th>
<th>Ionic species (in H₂O)</th>
<th>Limiting ionic conductivity (10⁻⁴ Ω⁻¹ mol⁻¹ m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 M KOH</td>
<td>20.1</td>
<td>5.0</td>
<td>25</td>
<td>20</td>
<td>H⁺</td>
<td>349.8</td>
</tr>
<tr>
<td>0.5 M KOH</td>
<td>10.7</td>
<td>9.3</td>
<td>47</td>
<td>18</td>
<td>K⁺</td>
<td>73.5</td>
</tr>
<tr>
<td>0.1 M KOH</td>
<td>2.26</td>
<td>44</td>
<td>221</td>
<td>18</td>
<td>Na⁺</td>
<td>50.1</td>
</tr>
<tr>
<td>0.5 M NaOH</td>
<td>8.6</td>
<td>12</td>
<td>58</td>
<td>18</td>
<td>OH⁻</td>
<td>197</td>
</tr>
<tr>
<td>Satd. NaCl (~5 M)</td>
<td>21.4</td>
<td>4.7</td>
<td>23</td>
<td>18</td>
<td>Cl⁻</td>
<td>76.4</td>
</tr>
<tr>
<td>1.0 M NaCl</td>
<td>7.44</td>
<td>13</td>
<td>67</td>
<td>18</td>
<td>SO₄²⁻</td>
<td>162</td>
</tr>
<tr>
<td>0.5 M NaCl</td>
<td>3.8</td>
<td>26</td>
<td>132</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1 M NaCl</td>
<td>1.07</td>
<td>93</td>
<td>467</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.5 M H₂SO₄</td>
<td>73.9</td>
<td>1.4</td>
<td>7</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0 M H₂SO₄</td>
<td>36.6</td>
<td>2.7</td>
<td>14</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5 M K₂SO₄</td>
<td>6.2</td>
<td>16</td>
<td>81</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distilled water</td>
<td>$10^{-3}$ to $10^{-4}$</td>
<td>$10^{3}$ to $10^{6}$</td>
<td>$\infty$</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Millipore water</td>
<td>$-5.5 \times 10^{-6}$</td>
<td>$-18 \times 10^{6}$</td>
<td>$\infty$</td>
<td>25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$R_E$ and $V_{loss}$ are calculated using $K_{cell} = 1$ cm⁻¹. Individual ionic conductivities are shown to give an impression of their relative contributions (note that listed values are for the limit of low concentrations, i.e., no interactions). Data compiled from refs. [16–18].
3.3 The Photoelectrochemical Measurements Setup

Photocurrent measurements are the main tool for investigating the properties and performance of photoelectrodes. Two main experimental configurations can be distinguished, based on the type of light source used: one for performance and stability measurements under high intensity white light, and the other for measuring wavelength-dependent properties using a (low-intensity) monochromatic light source. In this section, these two configurations and their components are discussed. Special attention will be given to the two main components: the potentiostat and the light source.

3.3.1 Simulated Sunlight Measurements

The ultimate test for any photoelectrochemical device or individual photoelectrode is its performance under solar irradiation. Since few places in the world have daily access to sunlight of a constant intensity, most PEC research laboratories are equipped with a solar simulator. Figure 3.4 shows an overview of the main components for a photoelectrochemical experimental test setup featuring simulated sunlight. A potentiostat is used to control the sample’s potential and to measure

![Diagram: Experimental setup for measuring the performance of a photoelectrode under irradiation with simulated sunlight](image)
the current, and an electromechanical shutter is used to block the light for, e.g., chopped-light experiments. A calibrated solar cell or photodiode is used to verify that the intensity of the solar simulator at the sample position corresponds to the desired value – usually 1,000 W/m², which corresponds to 1 sun at AM1.5. A more detailed description of the individual components will be given in the paragraphs below. The setup shown in Fig. 3.4 can be used for performance measurements as well as medium-term stability measurements. The latter are usually limited to 72 h; for longer term stability measurements dedicated testing facilities that also feature temperature and humidity control may be more suitable.

### 3.3.2 Wavelength-Dependent Measurements

More detailed insights into the factors that limit the performance of a photoelectrochemical device or photoelectrode can be obtained by measuring the photocurrent as a function of wavelength. A typical setup for such measurements is shown in Fig. 3.5. The light source is a tungsten halogen or gas discharge (e.g., xenon) lamp, and a monochromator is used to filter out a narrow part of the spectrum centered around the wavelength of interest. An electromechanical shutter is used to control the exposure of the sample to the light, and optical lenses are used to properly focus the light onto the monochromator and sample. If the dispersive element inside the monochromator is a diffraction grating, the monochromator output also contains
the higher orders of the diffracted light. These undesired components must be filtered out by the so-called long-pass filters (see below). The potential of the sample is again controlled by a potentiostat, which also measures the current. The intensity of the incident light is measured with a calibrated photodiode. In contrast to the reference cell in Fig. 3.4, this photodiode is specifically calibrated for – and meant to be used with – monochromatic irradiation. Using an internal calibration curve, it reports the incident light intensity for a single wavelength, either in Watts (=J/s) or in photons/s. From the photocurrent and the number of photons/s, one can readily calculate the external quantum efficiency or IPCE (cf. Sect. 2.8).

One aspect to be aware of with the setup of Fig. 3.5 is that the light intensities are 2–4 orders of magnitude lower than the 100 mW/cm² one gets with a solar simulator. Higher intensities of monochromatic illumination can be achieved by using a laser or one or more LEDs. These sources are convenient in the sense that long-pass filters are no longer necessary, and that the function of the expensive electromechanical shutter is often integrated in the LED driver. The main disadvantage, however, is that only one wavelength is available – perhaps a few more if one uses an argon ion laser or a series of different LEDs. One should also take note of the fact that the spectral half-width of LEDs is quite large, ranging between 10 and 50 nm (with 30 nm being a typical value).

### 3.3.3 The Potentiostat

A central component in every photoelectrochemical setup is the potentiostat. This is truly the work horse of photoelectrochemistry, but despite its important role it is often considered and treated as a black box. A basic understanding of its inner workings goes a long way in avoiding common user errors and getting the most out of one’s measurements.

A simplified diagram of a potentiostat is shown in Fig. 3.6. Most potentiostats have three or four test leads to be connected to the cell or device under test. The three leads that are always present are those for the working (WE), counter (CE), and reference (REF) electrodes. The optional fourth one is basically a second reference electrode which “senses” the potential of the working electrode. Depending on the manufacturer, it is labeled as “SENSE” or “RE2” or “REF2.” If no fourth test lead is present, it means that it is internally connected to the working electrode connection. The potentiostat’s aim in life is simple: it measures the potential difference between the REF and SENSE inputs, and makes sure that it stays equal to the desired potential difference by adjusting the potential at the CE output. The desired potential

---

7 For a more detailed description of the working principle of the differential amplifiers, buffers, and power amplifiers shown in Fig. 3.6, the reader is referred to the excellent classic text on electronics by Horowitz and Hill [23].
difference is the sum of the potential set by the user (either via the computer or by the potential control knob on the instrument) and the potential present at the modulation input.\(^8\)

The potential reported by the potentiostat corresponds to the measured potential difference between the REF and SENSE electrodes. So for a simple two-electrode measurement on a practical PEC device, the REF and CE leads have to be tied together in order to measure the potential difference between WE and CE. For a three-electrode measurement, the potential of the counter electrode is unknown. In most cases this not a problem, since we are only interested in processes that take place at the working electrode. The counter electrode only serves to supply the voltage and current that is necessary to keep \(V_{\text{REF-SENSE}}\) at the desired value.

The REF and SENSE inputs draw negligible current, so the entire current through the cell goes through the current range resistor, \(R_i\), which converts the current to a voltage. Most potentiostats have a \(R_i\) value of 10 \(\Omega\) for a 100 mA current range, 100 \(\Omega\) for a 10 mA range, etc., which results in a potential difference of 1 V across \(R_i\) when the current is at the maximum of the chosen range. Parallel to this resistor is a capacitor, \(C_{\text{filter}}\), whose value can often be selected (or altogether disconnected) by the user. Together with \(R_i\) this forms an RC filter that reduces the noise and gives more stable measurements. Most manufacturers specify the values of \(R_i\) and \(C_{\text{filter}}\) because their values affect the speed of the potentiostat – more on this in Sect. 3.6.5.

Standard photoelectrochemical measurements are a fairly undemanding application for modern potentiostats, and almost any commercial potentiostat will serve its purpose. It is nevertheless instructive to have a brief look at the main

---

\(^8\)The modulation input is used for, e.g., impedance measurements, which will be discussed later.
specifications and features. Table 3.4 shows some of the specifications for an EG&G PAR 283, which is a ~15-year-old research-grade potentiostat used in the author’s laboratory. Its specifications are still respectable – even by today’s standards – and it is more than able to handle the most demanding measurements for solar water splitting applications. The 283 has been succeeded by newer models and is no longer in production, so it simply serves to illustrate our discussion on specifications.

• The **maximum current** is one of the first specifications to consider when purchasing a new potentiostat. In practice, maximum current of 20–30 mA is more than sufficient so long as the electrode area does not exceed a few cm². Many potentiostats can deliver currents of 1 or 2 A, but this is overkill for most PEC research efforts.

  The **compliance voltage** is the maximum voltage that the power amplifier can deliver to the counter electrode. For poorly conducting electrolytes (e.g., organic solvents or ceramics) a high CE voltage may sometimes be necessary in order to achieve the required potential difference between the REF and SENSE inputs. For the highly conducting aqueous electrolytes in a PEC cell, however, this is

---

**Table 3.4** Specifications for a somewhat older type of research-grade potentiostat

<table>
<thead>
<tr>
<th>Specifications</th>
<th>EG&amp;G PAR 283 Potentiostat</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Power amplifier</strong></td>
<td></td>
</tr>
<tr>
<td>Max. current</td>
<td>±100 mA</td>
</tr>
<tr>
<td>Compliance voltage</td>
<td>±80 V</td>
</tr>
<tr>
<td>Slew rate</td>
<td>10 V/μs (no load)</td>
</tr>
<tr>
<td><strong>Electrometer</strong></td>
<td></td>
</tr>
<tr>
<td>Input bias current</td>
<td>&lt;10 pA</td>
</tr>
<tr>
<td>Band width (−3 dB)</td>
<td>&gt;9 MHz at $R_s = 1$ kΩ</td>
</tr>
<tr>
<td>Input impedance</td>
<td>&gt;10$^{13}$ Ω</td>
</tr>
<tr>
<td>Input capacitance</td>
<td>&lt;5 pF</td>
</tr>
<tr>
<td><strong>I and V measurement</strong></td>
<td></td>
</tr>
<tr>
<td>Bias potential resolution</td>
<td>14 bits, ±8 V</td>
</tr>
<tr>
<td>Lowest current range</td>
<td>1 nA</td>
</tr>
<tr>
<td>Current resolution</td>
<td>12 bits</td>
</tr>
<tr>
<td>Current accuracy</td>
<td>0.2% of full scale (2% for 1–10 nA current ranges)</td>
</tr>
<tr>
<td><strong>Frequency response</strong></td>
<td></td>
</tr>
<tr>
<td>−3 dB point</td>
<td>&gt;2 MHz @ 1 mA current range and 1 kΩ source impedance</td>
</tr>
<tr>
<td></td>
<td>&gt;200 kHz @ 1 μA current range and 1 MΩ source impedance</td>
</tr>
<tr>
<td><strong>Analog interface</strong></td>
<td></td>
</tr>
<tr>
<td>Analog I and V outputs</td>
<td>BNC, max. current corresponds to ±1 V, 50 Ω</td>
</tr>
<tr>
<td>External input</td>
<td>×1 (Ext. in) or × 0.02/×0.2</td>
</tr>
<tr>
<td><strong>Digital interface</strong></td>
<td></td>
</tr>
<tr>
<td>Aux. input/output</td>
<td>1 DAC out, ±10 V, 12 bits</td>
</tr>
<tr>
<td>Computer interface</td>
<td>GPIB, RS-232</td>
</tr>
<tr>
<td>Software/drivers</td>
<td>Dedicated software from manufacturer, Labview drivers available on the internet</td>
</tr>
</tbody>
</table>
never an issue and a compliance voltage as small as ±8 V is more than sufficient. In fact, a high compliance voltage can pose a serious safety problem. For example, if one accidentally disconnects the reference electrode during operation, the potential difference between REF and SENSE is undefined and the power amplifier may put the maximum compliance voltage on the counter electrode in order to compensate. Any voltage higher than 42 V (dc) is potentially lethal, so this is something to consider carefully, especially when the equipment is operated by individuals with limited experience in this field (students!).

- The **slew rate** indicates how fast the potential of the counter electrode can change to a different value. For standard voltammetry this value is not so relevant, but it may become important for, e.g., time-resolved photocurrent spectroscopy or impedance spectroscopy. Anything above 10 V/μs should be fast enough for PEC work.

- The **electrometer** refers to the differential amplifier that measures the REF and SENSE inputs. It should be able to do this while drawing negligible current from the reference electrode. The 283’s 5 pA input current is certainly small enough – in fact, anything below a few nA is acceptable. The same goes for the input impedance, anything above $10^{10}$ Ω would do. The **band width** determines how fast the electrometer is. Note that the 9 MHz in Table 3.4 does not mean one can do impedance measurements up to this frequency: this value refers to the so-called “-3 dB” point, which is the frequency at which the reported signal is only ~70% of the actual signal and shifted 45° in phase. The frequency should be at least 10× lower to get reasonable accuracies: at 0.9 MHz the amplitude difference is only 0.5%, and the phase shift is reduced to ~6°. The input capacitance is also important, since it forms a low-pass filter with the reference electrode’s impedance – more on this in Sect. 3.6.5.

- A **bias potential resolution** of 14 bits over a range of ±8 V means that the smallest potential step is $(8-(-8))/2^{14} \approx 1$ mV. This is certainly acceptable, although smaller values would allow even smoother voltammograms. The current resolution is where the 283 starts to show its age. To illustrate this, a 12-bit resolution for a ±2 mA current range means that the resolution is ~1 μA. Since dark currents are often more than $10^3$ times smaller than the maximum photocurrents, measuring both in a single chopped-illumination experiment requires switching the current range every time the light is switched on or off. This often leads to spikes in the measured current for capacitive samples and may even cause some of the faster potentiostats to start oscillating. Such oscillations should be avoided at all cost since they may destroy the sample. Newer potentiostats have 16-bit resolution D/A and A/D converters that allow very smooth voltammograms and less current range switching.

- The **frequency response** is not so relevant for voltammetry, but it is a crucially important figure-of-merit for impedance measurements. For now, suffice it to say that practical frequencies should be 10–50× smaller than the reported -3 dB frequency at the current range used. More details on impedance measurements are discussed in Sect. 3.6.
• The analog outputs of the cell voltage and the cell current are essential for impedance measurements using an external FRA, or for measuring low-level photocurrents using an external lock-in amplifier. The external input (sometimes called the modulation or polarization input) is used to super-impose a small sinusoidal voltage on the cell during impedance measurements. It can also be used to apply the large-amplitude modulation required for, e.g., linear sweep voltammetry, although the internal sweep generators inside modern potentiostats are so good that this is almost never necessary. All potentiostats have these very convenient analog inputs and outputs, except for some models that have a built-in impedance analyzer.

• Digital interface: some potentiostats feature auxiliary A/D inputs and D/A outputs that can be controlled by the software. These are convenient for controlling auxiliary equipment, such as electromechanical shutters or magnetic stirrers. GPIB used to be the standard interface to the computer for instrument control and data exchange. However, this >30-year-old standard is now outdated and is superseded by USB and/or Ethernet interfaces. The same goes for the even older – and much slower – RS-232 interface.

• The final item is the software that controls the potentiostat and is used to collect and transfer the data to a computer. The importance of a user-friendly piece of software to do daily measurements cannot be overstated. Fortunately, the software packages of most manufacturers have matured into powerful and easy-to-use measurement and analysis packages. While great for standard electrochemical measurements, it should be mentioned that they offer (very) limited built-in support for controlling additional instruments, such as monochromators, shutter drivers, calibrated photodiodes, and lock-in amplifiers. To get around this issue, the user has to write his own software. Today’s de facto standard for doing this is Labview®, a graphical programming language developed by National Instruments. Most – but not all – potentiostat manufacturers offer Labview drivers. If not, suitable drivers may available on the internet. Some manufacturers are unwilling they provide the necessary command set information that allows researchers to write their own control software – if this is the case, one should seriously consider another brand of potentiostat for PEC studies.

When starting to work with a new potentiostat, one should be aware that there exist different sign conventions for the reported voltage and current. To avoid confusion, we strongly recommend that one adheres to the following convention for photoelectrochemistry: (a) the reported potential should reflect the potential of the working electrode with respect to that of the reference electrode, and (b) more positive potentials should give more positive currents. In other words, more positive potentials favor (anodic) oxidation of the working electrode, and anodic currents

---

9 Most proprietary software packages include support for several types of frequency response analyzers, sometimes even for models made by other manufacturers.
are positive. Certain potentiostats do not follow this convention, and this can have the strange and confusing consequence that a two-electrode measurement on a simple resistor gives a negative current when the bias potential is positive. The sign convention used by a particular potentiostat can be easily determined by connecting a multimeter across the resistor in such a two-electrode measurement. We finish this paragraph with a list of well-known potentiostat manufacturers: Princeton Applied Research (PAR), Solartron Analytical, Metrohm Autolab, Ivium, Gamry, CH Instruments, and Bio-logic – and there are many others.

### 3.3.4 The Solar Simulator

The main figure-of-merit for a solar water splitting device is its performance under real sunlight. However, even real sunlight does not have the same intensity and spectral distribution everywhere on earth. To facilitate meaningful comparisons of device performances, the performance characteristics are usually quoted for the so-called **AM1.5G** conditions. This stands for “air mass 1.5 global,” and refers to the spectral distribution and intensity of sunlight on a 37° south-facing tilted surface after it has traveled through 1.5 times the thickness of the earth’s atmosphere. The latter corresponds to a solar zenith angle of 48.19°. The AM1.5G spectrum includes both the direct and the diffuse contributions of the incident sunlight, and has a total integrated intensity of 1,000 W/m².\(^{10}\) The most recent and widely used AM1.5G reference spectrum is published by the American Society for Testing and Materials as the ASTM-G173-03 standard, and represents a reasonable average for the 48 contiguous states of the USA over a period of 1 year [24]. The corresponding international standard is IEC 60904-3 (2nd edition, 2008).

Since real AM1.5G sunlight is not readily available at all times and at all locations, solar simulators are used. Making a reasonably good solar simulator is far from trivial. One cannot use halogen lamps since their maximum color temperature is limited to ~3,200 K, which is much lower than the color temperature of the sun (5,800 K). Xenon lamps are widely accepted to give the best match to the solar spectrum. Optical filters are used to optimize the overall spectral shape and to remove (to some extent) the sharp lines in the emission spectrum of xenon. A water filter is sometimes used to remove excess spectral intensity in the IR region. Despite such efforts, many conversion efficiencies of photoelectrochemical cells reported in the literature using primitive home-made solar simulators based on, e.g., Xe lamps are overestimated, as demonstrated in an important and carefully conducted study by Murphy et al. [25]. In order to obtain reliable efficiency estimates, Murphy et al. recommend two approaches: (a) measure the performance using actual sunlight under a similar zenith angle as that of the AM1.5 spectrum, or (a) measure the

---

\(^{10}\) The AM1.5D spectrum only contains the *direct* contributions over a 5° field of view, and has an integrated intensity of 768 W/m².
IPCE as a function of wavelength and integrate this over the AM1.5 solar spectrum (using (2.74) in Sect. 2.8). Unfortunately, the first approach is not always possible in certain parts of the world. The second approach should be treated with some caution: the intensity used during monochromatic quantum efficiency measurements is usually much lower than 1,000 W/m², and this can give large errors for samples that show a nonlinear photocurrent response (e.g., BiVO₄). To avoid this, one has to either verify the linearity of the photocurrent response up to 1,000 W/m² (using a laser), or try to measure the IPCE under simultaneous high-intensity background illumination.

An alternative option is to purchase a good quality commercial solar simulator. Various standards have been defined to evaluate the performance of solar simulators. The three main criteria are spectral matching, uniformity of irradiation, and temporal stability. The quality with regard to each of these criteria is indicated by three consecutive letters. Table 3.5 shows the widely adopted American, European, and Japanese standards for these criteria. If a manufacturer conforms to all three standards, it is the strictest one that counts. For example, an “ABB class” solar simulator that conforms to all three standards of Table 3.5 would have a spectral irradiance (W/m²/nm) that deviates no more than −25% to +25% from the AM1.5G intensity within the entire wavelength range (usually 400–1,100 nm). The nonuniformity of the integrated irradiation intensity (W/m²) should be within 3% for every point inside the specified area (e.g., 2 × 2 in.), and the integrated irradiation intensity would vary less than 1% (rms) in time. A 2 × 2 in. class ABB solar simulator would do very nicely for almost all PEC studies.

From Table 3.5, it is clear that the largest uncertainties are in the spectral matching criterion. To get a feeling for the extent of the deviations, Fig. 3.7
compares the spectrum of a commercial solar simulator to the AM1.5G and AM1.5D reference spectra. The intense xenon emission peaks around ~764, ~825, and above 880 nm can be safely ignored for PEC cells based on metal oxides, since their bandgap is usually larger than 1.7 eV (<730 nm). It is important to be aware of the series of peaks around 475 nm, even though the region between 450 and 500 nm represents only ~17% of the total spectral irradiance of a solar simulator between 300 and 1,100 nm for the solar simulator of Fig. 3.7, the match in this region to the AM1.5G and AM1.5D spectra is quite good. However, this should be considered an exception, and many solar simulators on the market show much poorer matching. One reason for this is that the international performance standards are only specified above 400 nm. This means that the spectral irradiance of a class ABB simulator may deviate much more than 25% below 400 nm, while still being “class A” in terms of spectral matching. The only
way to make sure that a solar simulator is indeed suited for metal oxide PEC research is by inspection of the solar irradiance curves from the manufacturer, or to measure this yourself with a calibrated spectroradiometer. The latter are sold by several suppliers, e.g., International Light Technologies or Ocean Optics.

When choosing a solar simulator for PEC research, one should carefully consider the orientation of the light exit port. Most solar simulators have their exit port oriented vertically, i.e., shining “down.” This is convenient for solid-state solar cells, but a bit of a hassle for PEC cells containing a liquid electrolyte – one would have to use a mirror since the photoelectrode is mounted vertically in most PEC cells. Fortunately, several manufacturers offer the possibility to have the light exit sideways. Some solar simulator manufacturers even offer integrated electromechanical shutters, with minimum exposure times as short as 200 ms and lifetimes $>10^6$ cycles. There are many manufacturers of solar simulators, and some well-known names include Oriel Instruments (Newport), Asahi, Abet Technologies, Luzchem, Wacom, Spire, Voss, etc.

### 3.3.5 Miscellaneous Components

As shown in Figs. 3.4 and 3.5, a photoelectrochemical setup contains several other components in addition to the potentiostat and solar simulator. We briefly discuss them below.

**Lamp** – For wavelength-dependent photocurrent measurements, a combination of a lamp and a monochromator is commonly used as a monochromatic light source (Fig. 3.5). The lamp should have a broad and homogeneous emission spectrum to ensure reasonable light intensities from the UV up to the near-IR. Xenon and tungsten-halogen lamps are most often used. Xe lamps have higher intensities, especially in the UV, and can be nicely collimated due to their small spot size. The downside is that Xe contains sharp emission lines around 475, 765, 825, and 880 nm, and that the short-term temporal stability is significantly worse than that of halogen lamps. Xe lamps are gas-discharge lamps and have to be ignited, which can cause large electromagnetic spikes during ignition. Without proper shielding by the manufacturer, this can interfere with sensitive electronic equipment and bring a potentiostat into oscillation, possibly destroying the sample.\(^{11}\) This is not a problem for tungsten–halogen lamps. These lamps give very smooth spectra without specific emission peaks, and still give reasonable output power above 300 nm when a “quartz” (i.e., fused silica) lamp envelope is used. The tungsten filament is usually coiled and has a certain size, typically $7 \times 4$ mm\(^2\) for a 250 W lamp. The highest light intensity is achieved when the projected image of the filament exactly fills the filament envelope.

\(^{11}\) In exceptional cases, ignition of a high-power gas discharge lamp can even destroy electronic equipment. This has happened once in the author’s laboratory to a Solartron 1286 potentiostat after switching on an older type 450 W Xe lamp.
Choosing a lamp with a higher power rating does not give higher light intensities at the sample, since the filament temperature is limited to ~3,200 K. Despite their somewhat lower irradiation intensity, quartz tungsten-halogen lamps are often preferred for PEC research due to their low cost, temporal stability, and ease of use. Manufacturers of Xe and halogen light sources include Oriel (Newport), OBB Corp, Sutter, Cermax, Hamamatsu, PTI, etc.

**Monochromator** – The monochromator filters the desired wavelength out of the incident broad-band irradiation. The dispersive element is usually a grating, which consists of a series of closely spaced sawtooth-shaped grooves that are etched in a highly reflecting mirror surface. The desired output wavelength is selected by rotating the grating with a stepper motor. Figure 3.8 shows a simplified diagram of a Czerny–Turner monochromator, one of the most common designs. For PEC research, one of the main selection criteria is to get as much light as possible into the monochromator. This can be achieved by selecting a modest focal length (0.1–0.3 m) and the largest diameter grating one can afford. This translates into a small $F$-number ($F/$#), which is defined as the focal length divided by the aperture (diameter) of the optical element. For example, a 150 mm monochromator with a 36 mm grating has an $F$-number of $F/4$. An $F/2$ monochromator has more light gathering power, but will also be more expensive due to the larger grating and mirrors. To ensure optimal filling of the grating with the incident light, the lens in front of the monochromator entrance should have a matching $F/$#. A lens with a larger $F/$# does not fully illuminate the grating and will lead to lower intensities at the output. A smaller $F/$# may overfill the grating and lead to stray light inside the monochromator. The latter should be avoided to prevent undesired wavelength components reaching the output. An aperture stop (variable-size aperture) in between the lamp and the lens is sometimes used to avoid overfilling of the grating (Fig. 3.8). A very convenient tool to check for stray light at the output is a miniature spectrometer. These spectrometers, offered by, e.g., Ocean Optics, have an optical...
fiber input and contain a grating and a CCD camera in a very small package that can be directly powered by a USB port. The easiest way to ensure optimal illumination of the monochromator entrance and minimize stray light is by purchasing light sources with outputs that are specifically matched to monochromator inputs (sometimes called “monochromator illuminators”). These are available from, e.g., Princeton Instruments and Oriel/Newport.

At the output side of the monochromator, the light exits through a slit. The width of the slit is variable (usually between 10 μm and 3 mm) and determines the intensity and wavelength range of the light. An important parameter in this regard is the linear dispersion of the monochromator. Its value can be estimated with the expression

\[ \frac{d\lambda}{dx} = \frac{10^6 \cos \beta}{Nkf} \tag{3.3} \]

where \( \beta \) is the angle between the grating normal and the diffracted light, \( m \) is the diffraction order (monochromators always use the first order, so \( m = 1 \)), \( N \) is the number of grooves/mm of the grating, and \( f \) is the effective focal length in mm. The value of \( \beta \) depends on the wavelength and varies between \( \sim 17^\circ \) and \( 35^\circ \). As an example, consider a \( f = 150 \) mm monochromator with a 600 g/mm grating. Equation (3.3) predicts a linear dispersion of \( \sim 10 \) nm/mm. Even a slit opening as large as 1 mm\(^2\) (to obtain sufficient output intensity) would still give sufficient optical resolution for PEC studies. A 1200 g/mm grating would improve the resolution by a factor of two, but also results in a narrower wavelength range. In practice, the wavelength range is limited by the efficiency of the grating (i.e., not by the mechanical range of the monochromator). A wavelength range between 300 and 800 nm is well suited for PEC work, and most manufacturers offer gratings that cover this range. It is important to realize that grating monochromators also transmit higher order diffracted light. So when the monochromator is set to 600 nm, it also contains 300, 200, and 150 nm components that need to be filtered out (see below). Again, a miniature spectrometer is a convenient tool to quickly check the presence of second order components, and also to verify the spectral width of the emitted light. Gratings can be *blazed*, which means that the angles of the grooves are optimized to give the highest reflection at a specific wavelength. As a rule of thumb, gratings can be used between 2/3 and 2 times the blaze wavelength. So for most PEC work, a grating blazed at 400 or 500 nm would be optimal. For more detailed information on monochromators and gratings, the reader is referred to general optics textbooks [26–28] and the extensive information available on several manufacturer’s websites [29–31].

**Lenses** – The quality of the lenses does not need to be as good as, e.g., best form laser lenses, but they do need to be made from fused silica to ensure transparency at UV wavelengths. Lens diameters of 50 mm (2 in.) are generally sufficient to capture all the light from the lamp and from the monochromator exit.

\(^{12}\) Monochromators have both entrance and exit slits. It is recommended to use the same width for both slits.
Long pass filters (sometimes called high pass filters) only transmit light above a certain wavelength. They are essential for removing higher order diffracted light from grating-type monochromators. They are available as colored glass filters that are either square or round, with sizes ranging from 12 to 50 mm and a thickness between 1 and 3 mm. The type of filter is indicated as, e.g., WG280 or GG495, where the number represents the wavelength at which the transmission is 50% ($\lambda_{50\%}$). The transition is very steep: for a 3 mm filter, the transmission can change from $<1\%$ to $>85\%$ within 20–30 nm. The main manufacturer of these filters, Schott AG, offers an Excel-based tool that describes the spectral characteristics in detail [32]. Another manufacturer of long pass filters is Hoya.

Electromechanical shutters are used to quickly switch the illumination on and off. They are offered with diameters ranging between 1 and 90 mm, and with switching times down to a few milliseconds for the smaller models. The electromagnetic coils are driven by dedicated shutter drivers that have to be purchased separately. Depending on the model, these drivers can be controlled externally via a TTL-compatible input and/or they have their own timing electronics, usually with a TTL-compatible sync output. The shutter blades are often coated with a PTFE (Teflon®) coating to minimize the resistance. To avoid excessive heating of the shutter blades when using lasers or other high-intensity sources, certain models can be ordered with a reflecting metallic coating on one side of the shutter blades. For setups with a monochromator, a small-size shutter with an aperture somewhat bigger than the slit size can be placed directly after the monochromator (Fig. 3.5). Alternatively, the shutter can be placed in between the lens and the sample, at the focus point of the lens where the spot is at its smallest size. A shutter aperture of 5–6 mm in diameter is a suitable choice in most cases. Research-grade shutters of this size can have opening times as short as ~1–2 ms and continuous operating frequencies as high as 20 Hz. For setups with a solar simulator, shutters with a larger area are usually required to ensure full illumination of the sample. As mentioned in Sect. 3.3.4, some solar simulator manufacturers offer internal shutters with opening times of ~200 ms. If this is too slow, research-grade shutters with a clear aperture of 45 mm are available with opening times as small as 14 ms and continuous repetition rates of 2.5 Hz. Vincent Associates (Uniblitz® shutters) and CVI Melles Griot are well-known manufacturers of electromechanical shutter systems.

Calibrated photodiodes fall in two categories: those used to calibrate white light sources (e.g., solar simulators) and those calibrated for measuring one wavelength at a time for a certain range of wavelengths. The output of a solar simulator is usually measured with a calibrated PV solar cell or Si photodiode. An important – but often overlooked – assumption is that the spectrum of the solar simulator corresponds to the spectrum used during calibration. Because of this, it is probably best to purchase the photodiode and solar simulator from the same manufacturer. Alternatively, one can use a thermopile for calibrating white light sources. A thermopile contains a black body material that converts the incident radiation into heat, and the temperature is measured by a large number of thermocouples connected in series. It gives an output voltage that is proportional to the absolute
power of the incident light (in W or in W/m²) and its response covers a wide range of wavelengths – depending on the type from the infrared (> 5 μm) down to the UV (200 nm). It is almost insensitive to the spectral shape of the light source, but it has a much slower response time (>1 s) than a Si-based detector.

The intensity of monochromatic light sources can be determined with calibrated Si photodiodes that come with a wavelength-dependent response curve. The response is defined in units of A/W. Calibrated photodiodes are available from many manufacturers, but finding types that can be used in the UV (down to 300 nm for metal oxide PEC research) can be challenging. The PD300-UV from Ophir Optics is a convenient model used in the author’s laboratory which covers a wavelength range of 200–1,100 nm and powers ranging from 20 pW to 300 mW. The calibration curve is stored within a separate controller, which is connected to the measurement computer via a USB or RS-232 interface. The program sends the current wavelength to the controller, after which the measured output (in Watts) can be read back.

3.4 Photocurrent–Voltage Measurements

Current–voltage measurements are the most important technique to determine the performance characteristics of photoanodes or photocathodes for water splitting. It allows one to quickly obtain the following information:

- The photocurrent density and energy conversion efficiency as a function of applied potential.
- The dark current as a function of the applied potential.
- The sign of the photocurrent (anodic or cathodic).
- The photocurrent onset potential.
- Electron- vs. hole-transport limitations (front- and back-side illumination).
- Transient effects that indicate recombination.

The measurements are done by measuring the (photo)current while sweeping the potential at a constant scan rate – usually between 1 and 50 mV/s. This can be done in the dark, under constant illumination, or by switching the illumination on and off at a fixed frequency.

3.4.1 Dark Current and Photocurrent Onset Potential

Figure 3.9a shows a linear sweep voltammogram of a 0.2% Si-doped α-Fe₂O₃ photoanode deposited on FTO glass. At potentials more positive than ~1.7 V_{RHE}, the dark current starts to increase, indicating the electrochemical oxidation of water at the surface. This is due to an increase in the concentration of free holes in the valence band as the Fermi level is pulled down (more positive potential) and starts to approach the valence band in the region near the surface. A positive photocurrent is observed upon illumination, as indeed expected for an n-type photoanode.
The photocurrent onset potential is at ~0.95 V vs. RHE. This is much more positive than the flatband of α-Fe₂O₃, which is at ~0.3 VRHE [34, 35]. The difference is due to slow oxidation kinetics (catalysis) [36] and/or to recombination in the space charge region – see Sect. 2.6.6.

3.4.2 “Water Splitting Efficiencies”

If the incident light intensity is known, the energy conversion efficiency can be calculated from the photocurrent density. If one assumes that 100% of the photocurrent is used to actually split water (a reasonable assumption for a concentrated KOH or NaOH electrolyte), this also gives the solar-to-hydrogen conversion efficiency. The question is then at which applied bias potential one should take the photocurrent. Ideally, this should be at zero bias for a two-electrode measurement, and presumably also at zero bias vs. RHE for a three-electrode measurement. However, at these bias potentials the photocurrent is often very

---

13 The situation is a bit complicated in a three-electrode system since the potential of the counter electrode is not known. One can argue that for a counter-electrode with a low overpotential for H₂ evolution, the potential of the counter electrode should be close to that of the reversible hydrogen potential – provided that the concentration of dissolved hydrogen is high enough. This can be easily verified by a separate measurement of \( V_{WE} - V_{CE} \) during a three-electrode linear sweep voltammogram. One particular test in the author’s laboratory showed that at a current density >0.5 mA/cm², the potential of a coiled Pt wire electrode approached that of the reversible hydrogen electrode to within ~0.1 V. At lower current densities, however, deviations as large as 1 V were observed.
small for a cell with a single photoelectrode (see, e.g., Fig. 3.9) due to unsuitable band edge positions or low overpotentials/slow catalysis. To allow a meaningful performance comparison between different photoelectrodes reported in the literature, the photocurrent densities at 1.23 \textit{V}_{\text{RHE}} (for photoanodes) and 0 \textit{V}_{\text{RHE}} (for photocathodes) are often used as performance metrics. However, it is crucial to realize that the efficiency for an individual photoanode is actually zero for a bias of 1.23 \textit{V}_{\text{RHE}}, as can be seen from the expression for the solar-to-hydrogen efficiency:

$$\eta_{\text{STH}} = \frac{j_{\text{photo}}(V_{\text{redox}} - V_{\text{bias}})}{P_{\text{light}}}. \quad (3.4)$$

This expression is the so-called “applied bias photon-to-current efficiency,” or \textit{ABPE} [1]. Any claim of a certain solar-to-hydrogen efficiency for a single photoanode at a nonzero bias potential should therefore be understood as an extrapolation toward a tandem cell device, in which the required bias voltage is generated by a photovoltaic cell that is placed \textit{behind} the photoanode [37, 38]. This PV cell should of course have an equally large area as the photoanode, so that it only absorbs that part of the solar spectrum which is not absorbed by the photoanode. This important distinction is often overlooked in the literature. Moreover, it should be realized that efficient tandem architectures can be difficult to achieve in practice. This is particularly true for nanoporous photoelectrodes, in which extensive scattering may prevent the light from actually reaching the PV device [39]. For these reasons, claims of high “water splitting efficiencies” for any photoelectrode that requires a bias potential should be strongly discouraged. The only way in which a realistic value for the water splitting efficiency can be obtained is to measure the current for a two-electrode device under short circuit conditions.

### 3.4.3 Front- and Back-Side Illumination

An interesting observation in Fig. 3.9a is that the photocurrent under front-side illumination (via the electrolyte) is much larger than that under back-side illumination (via the substrate). As first pointed out by Lindquist et al. [40], a dependence of the photocurrent on the illumination direction indicates either electron or hole transport limitations in the bulk of the film. This is illustrated in Fig. 3.9b. Under front-side illumination, most light is absorbed near the semiconductor/electrolyte interface. This means that the photogenerated electrons have to travel a larger distance before reaching the interface than the photogenerated holes. For back-side illumination, the situation is reversed. Since surface kinetics, band bending, and other aspects are not (or hardly) affected by the direction of illumination, the smaller back-side illumination photocurrent observed in Fig. 3.9a implies that hole transport is more difficult than electron transport in Fe$_2$O$_3$. This is indeed
consistent with the very small hole diffusion length of 4–20 nm reported for hematite – see Chap. 4 for a more extensive discussion on this. For undoped Fe$_2$O$_3$ films the situation is reversed: the photocurrent under front-side illumination is lower (not shown), indicating that the material behaves as an insulator and that slow electron transport limits the photocurrent. Note that differences in front- and back-side illumination due to slow carrier transport can only be observed when a substantial light intensity gradient is present in the film. A good rule of thumb is that ($\alpha L$) should be $>1$, where $\alpha$ is the absorption coefficient and $L$ the film thickness.

3.4.4 Photocurrent Transients

The photocurrent–voltage curve is sometimes recorded while switching the light on and off with a shutter. An example of such a “chopped” AM1.5 voltammogram is shown in Fig. 3.10a for a BiVO$_4$ photoanode. Small differences between the dark current and the photocurrent can be easily observed with this technique, and this can help to accurately determine the photocurrent onset potential. Note that Fig. 3.10a shows pronounced transients when the light is switched on. Such transients are often observed and indicate recombination of the photogenerated charge carriers. To explain this, consider the transient photocurrent response shown in Fig. 3.10b for a BiVO$_4$ photoanode illuminated with 364 nm light from a continuous-wave Ar$^+$ laser (cw) at a fixed potential of 1.23 V$_{RHE}$ (the $x$-axis in this figure represents the time instead of the potential, but this does not affect the discussion). When the light is switched on, the photocurrent instantaneously rises to its peak value ($j_0$). After that, it decays to a steady-state value ($j_{SS}$) as part of the photogenerated electrons and holes ($j_0 - j_{SS}$) recombine. This recombination can be due to accumulation of holes near the surface, accumulation of electrons in the bulk (i.e., slow electron transport), or trapping of electrons or holes at surface states.
The accumulation of holes would cause an equally large cathodic transient when the light is turned off and electrons in the conduction band react with the accumulated holes. Only a small cathodic transient is observed in Fig. 3.10b, so accumulation of holes can be ruled out as the main recombination mechanism. A useful check for hole accumulation in a photoanode is by adding a hole scavenger, such as methanol [22] or hydrogen peroxide, or by catalyzing the water oxidation reaction with a co-catalyst such as IrO$_x$ [9, 36], Co$^{II/III}$ [21], RuO$_2$ [41, 42], or cobalt phosphate [43]. Recombination via surface states is also unable to explain the observed photocurrent transients. This is because the occupation of surface states is determined by the applied potential, and no transient is observed at the same potential but at lower light intensities (not shown). For an in-depth discussion on surface state recombination the reader is referred to two interesting papers by Peter [44] and Salvador [45], who pioneered the detailed analysis of photocurrent transients in the early 1980s. The origin of the transients in Fig. 3.10 is believed to be due to the poor electron transport properties of BiVO$_4$. This is indeed consistent with the observation that the photocurrent of BiVO$_4$ for front-side illumination is always lower than for back-side illumination.

3.5 Photocurrent Action Spectra

The spectral response of a photoelectrode can be investigated by measuring so-called photocurrent action spectra, in which the photocurrent is recorded as a function of the wavelength at a constant applied potential. One usually also records the intensity of the incident light as a function of wavelength in a separate measurement using a calibrated photodiode. The following information can be obtained from photocurrent action spectra:

• The incident photon-to-current conversion efficiency (IPCE).
• An indication in which part of the film more/less recombination occurs.
• An estimate of the photoelectrode performance under AM1.5 conditions.
• The sign of the photocurrent.
• Electron- vs. hole-transport limitations (front- and back-side illumination).
• Transient effects that indicate recombination.

While there is clearly some overlap between the information that can be obtained from a photocurrent voltammogram (Sect. 3.4) and an action spectrum, it should be realized that action spectra are generally recorded at much lower light intensities. This means that the measured currents under illumination are relatively small, and the dark current may no longer be negligible. The photocurrent therefore always has to be reported as the difference between the current under illumination and the dark current ($i_{\text{photo}} = i_{\text{light}} - i_{\text{dark}}$). Since most samples show some degree of current drift over a time scale of 5–10 min, $i_{\text{light}}$ and $i_{\text{dark}}$ are best recorded in a single measurement by chopping the light on and off. Care should be taken to ensure that
the current has reached a stable value before it is recorded. The preferred scan direction is from high wavelengths to low wavelengths, i.e., from low-energy photons to high-energy photons. This minimizes the chance that the sample itself changes during the measurement, e.g., through the creation of surface states during illumination. Photocurrent action spectra are usually recorded in two parts to allow changing the long pass filters that block the second order of the diffracted light from the monochromator. For example, a scan between 300 and 700 nm can be divided into a scan from 700 → 530 nm using a GG495 filter, and from 530 → 300 nm using a WG280 filter. If the region around 530 nm contains important information, one can also switch filters at shorter wavelengths (e.g., 430 nm using a GG395 filter) to avoid distracting discontinuities in the photocurrent.14

3.5.1 Quantum Efficiencies

The IPCE – also called the external quantum efficiency – is obtained by dividing the measured photocurrent by the number of photons using (2.71). The difference between the measured IPCE and the maximum value of 100% reflects the sum of the incident photons that are reflected or transmitted (instead of absorbed), and the number of photogenerated electron–hole pairs that recombine before reaching the outer circuit. The absorbed photon-to-current conversion efficiency (APCE) only reflects losses due to charge carrier recombination within the semiconductor, i.e., it is corrected for reflection and transmission losses. This is also referred to as the internal quantum efficiency. The APCE is a more useful parameter than the IPCE when trying to optimize the material in terms of purity, crystallinity, and defect density. It can be obtained from the IPCE by dividing it through the optical absorption of the sample using (2.72).

An example of an IPCE curve for a (rather poor) BiVO4 photoanode is shown in Fig. 3.11. Less than 10% of the incident photons end up contributing to the photocurrent. The other 90% is lost due to recombination and reflection + transmission of the incident light. Fig. 3.11 also shows that front-side illumination gives lower photocurrents than back-side illumination, consistent with the poor electron transport in BiVO4 (cf. Sect. 3.4.4). The decrease in IPCE below 330 nm for front-side illumination is consistent with electron transport being a limiting factor; at this wavelength the absorption coefficient is so high that all the light is absorbed near the semiconductor/electrolyte interface, and the electrons have to travel through the entire film to reach the back contact. The decrease below 350 nm for back-side illumination is due to the absorption of the FTO substrate.

14 In contrast to photocurrent measurements, discontinuities in the ICPE spectrum can be easily avoided by using the same combination of wavelength range and long pass filters while measuring the light intensity with the calibrated photodiode.
Several examples of IPCE and APCE spectra for another material, $\alpha$-Fe$_2$O$_3$ (hematite), are shown in Chap. 4. This chapter discusses the various features of these spectra in detail, and nicely illustrates how these spectra can help in determining the regions of the film in which recombination is most pronounced.

### 3.5.2 Extrapolation to AM1.5 Performance

From the IPCE spectrum, an estimate of the total photocurrent under AM1.5 illumination can be obtained by multiplying the IPCE value with the AM1.5 photon flux for each wavelength, followed by integration over all wavelengths. The expression for this total solar photocurrent is given by (2.74) in Sect. 2.8. The value obtained should match the current measured under actual AM1.5 conditions provided that the sample shows a linear response. As outlined in Sect. 2.8, the presence of surface traps or slow surface reaction kinetics may cause nonlinear behavior, and can result in large deviations between the predicted solar photocurrent and the actual value under AM1.5 conditions. The effect of slow surface reaction kinetics has recently been observed in the author’s laboratory for BiVO$_4$. The poor water oxidation kinetics for this (un-catalyzed) material under high light intensities results in an AM1.5 photocurrent that is $\sim$8 times smaller than that predicted from the IPCE measured at low light intensities using (2.74) [46].

### 3.6 Impedance and Mott–Schottky Measurements

Of the techniques discussed in this chapter, impedance measurements are arguably the most complicated measurements to perform and interpret. The discussion below is limited to the most common form of impedance analysis used in solar water splitting research: measuring a Mott–Schottky plot to determine the donor density.

---

**Fig. 3.11** IPCE spectrum for a spray-deposited BiVO$_4$ photoanode on FTO glass, recorded at a potential of 1.23 $V_{RHE}$ in a 0.15 M K$_2$SO$_4$ aqueous electrolyte solution.
and flatband potential of a thin film photoelectrode. The basic principles of the measurement are discussed, and the necessary precautions to ensure a valid data set and correct interpretation are outlined.

### 3.6.1 The Impedance of a Thin Film Photoelectrode

The donor density \(N_D\) in a semiconductor film can be determined by measuring the space charge capacitance, \(C_{SC}\), as a function of the applied potential, \(\phi_A\). Since the derivation of this relationship was already given in Sect. 2.5.2, we only repeat the final result, which is the Mott–Schottky equation:

\[
\frac{1}{C^2_{SC}} = \frac{2}{\varepsilon_0 \varepsilon_r e N_D A^2} \left( \phi_A - \phi_{FB} - \frac{kT}{e} \right).
\]

(3.5)

A plot of the slope of \(1/C_{SC}^2\) vs. \(\phi_A\) should give a straight line with a slope that is proportional to \(1/N_D\), while the intercept with the potential axis gives the flatband potential, \(\phi_{FB}\).

The challenge is to distinguish \(C_{SC}\) from other contributions of the photoelectrochemical cell. These can be either capacitive or resistive in nature.\(^{15}\) Other capacitive elements are the Helmholtz capacitance and surface state capacitances. Examples of resistive elements are the semiconductor bulk, the electrolyte, contacts, and charge transfer resistances. The contributions of all these elements can be modeled by a so-called equivalent circuit. Figure 3.12a describes most of the contributions that can be found in metal oxide semiconductor/electrolyte systems. Extracting the space charge capacitance from the overall response of such a complicated circuit can be difficult. Fortunately, many contributions can be ignored, and in the most simple case this leads to the circuit shown in Fig. 3.12b. The overall impedance of this simple \(RC\) series circuit is given by

\[
Z' = R - j \frac{1}{\omega C_{SC}}
\]

(3.6)

Here, \(\omega\) represent the angular frequency at which the impedance is measured. It is equal to \(2\pi f\), with \(f\) the frequency in Hz. The presence of \(j (-\sqrt{-1})\) in (3.6) indicates that \(Z'\) is a complex number – the asterisk serves as a reminder of this. It can

\(^{15}\)There are several other, more complicated elements available to describe the various processes that can occur in a photoelectrochemical cell, such as the Warburg element (to model diffusion), the Constant Phase Element (CPE, used to describe processes that have a distribution of time constants or activation energies), and transmission lines (to model porous electrodes [47]). Porous electrodes and CPE elements that represent nonideal capacitive elements are briefly discussed below. For more detailed information, the reader is referred to the literature [48, 49].
therefore be represented as a vector in the complex plane, as shown in Fig. 3.13a. For this simple \( RC \) series circuit, it can be easily seen that the real part of the impedance, \( Z_{Re} \), corresponds to the resistance \( R \), and the imaginary part, \( Z_{Im} \), is equal to \((-1/\omega C_{SC})\). This means that the space charge capacitance can be directly determined from the imaginary part of the impedance. Thus, by measuring the imaginary part of the impedance as a function of the applied potential, one can obtain a Mott–Schottky curve in which \( 1/C_{SC}^2 \) is plotted against \( \phi_A \).

### 3.6.2 The Mott–Schottky Plot: An Example

Before going into the details of the measurement principles, equipment, and precautions, we briefly illustrate a few basic features of a Mott–Schottky plot using the example shown in Fig. 3.14 [50]. The Mott–Schottky curves in this figure...
are for dense, electron-beam evaporated anatase TiO2 films with a thickness of 40, 80, and 120 nm. The material is n-type, which implies that the depletion layer width increases when the potential is made more positive. Extrapolation of the steep slope between \( \Phi_0 = 1.0 \) and \( \Phi_0 = 0.8 \) V vs. SCE gives the flatband potential of TiO2. The three curves give nearly identical values for \( f_{FB} \), with an average value of \((-0.98 \pm 0.04)\) V vs. SCE. From the slopes, donor densities between \(4.1 \times 10^{16}\) and \(2.9 \times 10^{17}\) cm\(^{-3}\) are found for the films.

An interesting feature of the Mott–Schottky curves in Fig. 3.14 is the capacitance plateau found at potentials positive of \(-0.7\) V vs. SCE. At these potentials, the depletion layer width actually exceeds the TiO2 film thickness, and further extends into the underlying (n-type) ITO film. Clearly, such a plateau will only be observed for low donor densities and/or very thin films. The slope of the curve at the plateau is determined by the donor density of the ITO, for which a value of \(1 \times 10^{20}\) cm\(^{-3}\) is found in this particular case [50, 51]. Moreover, from the capacitance \( C_L \) at the intersection of both curves (dashed lines in Fig. 3.14) and the known thickness \( L \) of the films, it is possible to determine the dielectric constant of the TiO2 using the expression for a parallel plate capacitor \( C_L = \varepsilon_0 \varepsilon_r A/L \). A value of 55 was found for the dielectric constant of polycrystalline anatase TiO2 (this value was also used to calculate the donor densities mentioned in the previous paragraph).

### 3.6.3 The Frequency Response Analyzer

The real and imaginary parts of the impedance of the sample can be measured with a FRA. This is a fairly complicated machine that can be purchased from several companies. Some are sold as stand-alone machines (Solartron, Princeton Applied
Research), while others are sold as an integral part of a potentiostat or as a modules that can be separately purchased (Princeton Applied Research, Gamry, Metrohm-Autolab, Ivium).

While the FRA is often treated as a black box by the photoelectrochemist, a basic understanding of its inner workings may help to avoid measurement errors. We briefly describe its operating principles using the diagram shown in Fig. 3.15. The FRA generates a small sinusoidal voltage $V_{\text{ref}}$ of a certain frequency ($\omega_{\text{ref}}$) and amplitude ($A_{\text{ref}}$), which is used to modulate the potential of the sample via the modulation input of the potentiostat. When the system behaves linearly, this results in a sinusoidal response of the current that only differs in amplitude and phase from the voltage perturbation (Fig. 3.13b). The measured potential and current are then fed back into the FRA via the analog outputs of the potentiostat, and are further processed inside the FRA in order to calculate the impedance.

To ensure a linear response of the system, the amplitude of the perturbation voltage needs to be small, usually 5–10 mV for photoelectrochemical devices. Such a small perturbation also gives a very small current response. The first job of the FRA is to isolate these signals from the background signal. The latter consists of the steady-state (dc) current and voltage plus the unavoidable electrical noise. This seems an impossible task, since the signals that we are interested in can easily

![Fig. 3.15 Schematic diagram of a frequency response analyzer (FRA), showing the proper connections to the potentiostat for impedance spectroscopy and Mott–Schottky measurements. The function of the quasi-reference electrode is explained in Sect. 3.6.5](image-url)
be $1,000 \times$ smaller than the background signals. Fortunately, the FRA can do this very well using a so-called phase sensitive detector (PSD). To understand how this works, consider a sinusoidal input signal that can be described by $V_{\text{sig}} = A_{\text{sig}} \sin(\omega_{\text{sig}} t + \theta_{\text{sig}})$. After some preconditioning, the PSD multiplies this signal with its internally generated reference signal, $V_{\text{ref}} = A_{\text{ref}} \sin(\omega_{\text{ref}} t + \theta_{\text{ref}})$. Some trigonometric manipulation shows that the resulting signal can be written as

$$V_{\text{mult}} = A_{\text{ref}} A_{\text{sig}} \sin(\omega_{\text{ref}} t + \theta_{\text{ref}}) \sin(\omega_{\text{sig}} t + \theta_{\text{sig}})$$

$$= \frac{1}{2} A_{\text{ref}} A_{\text{sig}} \cos[(\omega_{\text{ref}} - \omega_{\text{sig}}) t + \theta_{\text{ref}} - \theta_{\text{sig}}]$$

$$- \frac{1}{2} A_{\text{ref}} A_{\text{sig}} \cos[(\omega_{\text{ref}} + \omega_{\text{sig}}) t + \theta_{\text{ref}} + \theta_{\text{sig}}].$$

(3.7)

The resulting signal has two frequency components, one at $(\omega_{\text{ref}} + \omega_{\text{sig}})$ and one at $(\omega_{\text{ref}} - \omega_{\text{sig}})$. The PSD then proceeds to integrate this signal during an integer number of cycles of the reference signal. If the signal frequency is the same as the reference signal’s frequency $(\omega_{\text{ref}} = \omega_{\text{sig}})$, the PSD give a dc output voltage that is proportional to the signal’s amplitude:

$$V_{\text{PSD}} = \frac{1}{2} A_{\text{sig}} A_{\text{ref}} \cos(\theta_{\text{ref}} - \theta_{\text{sig}}).$$

(3.8)

More importantly, the result of this integration is zero for all other signal frequencies. This means that the PSD very effectively rejects noise and other undesired signal components; it only detects the signal of interest, i.e., any signal that has the same frequency as the reference signal. The output of the PSD is further amplified and converted into a digital value by an A/D converter in the FRA.

As shown in Fig. 3.15, there are actually four phase-sensitive detectors in a FRA: two for the voltage and two for the current. Each signal is multiplied by the reference signal and by a signal that is shifted $90^\circ$ in phase from the reference signal. The latter is given by $A_{\text{ref}} \sin(\omega_{\text{ref}} t + \theta_{\text{ref}} + 90^\circ) = A_{\text{ref}} \cos(\omega_{\text{ref}} t + \theta_{\text{ref}})$. This results in two values that represent the real and imaginary components of each signal. To see this, remember that the voltage and current signals can also be written in complex notation:

$$V_{\text{sig}} = V_0 \exp(j(\omega_{\text{sig}} t + \theta_\text{V})) = V_0 \cos(\theta_\text{V}) + jV_0 \sin(\theta_\text{V}) = V_\text{Re} + jV_\text{Im},$$

$$I_{\text{sig}} = I_0 \exp(j(\omega_{\text{sig}} t + \theta_\text{I})) = I_0 \cos(\theta_\text{I}) + jI_0 \sin(\theta_\text{I}) = I_\text{Re} + jI_\text{Im}. (3.9)$$

The in-phase signal of the voltage is equal to $V_\text{Re}$, while the $90^\circ$ out-of-phase signal is the imaginary part, $V_\text{Im}$ (and similarly for the current). As a final step, the

16 Preconditioning steps are, e.g., removal of the dc component of the signal by passing it through a high pass filter (a capacitor), or amplification/attenuation of the signal by a certain fixed factor. These steps are often necessary to ensure that the signal falls within the range that the FRA can handle internally. These steps do not (and should not) influence the final measurement results.
microprocessor of the FRA divides the voltage by the current to give the complex impedance of the system in the form of numerical values for $Z_{\text{im}}$ and $Z_{\text{re}}$:

$$Z^* = \frac{V_{\text{Re}} + jV_{\text{Im}}}{I_{\text{Re}} + jI_{\text{Im}}} = Z_{\text{Re}} + jZ_{\text{Im}}.$$  

(3.10)

As shown in Sect. 3.6.1, we can now calculate the space charge capacitance by multiplying $Z_{\text{im}}$ with $-\omega$.  

### 3.6.4 Measurement Parameters

To measure a Mott–Schottky plot, the potential is swept at a constant scan rate while continuously measuring the imaginary part of the impedance at a fixed modulation frequency. The starting potential is a potential where a large depletion layer is present. The scan then proceeds in the direction of the flatband potential, and is reversed (or stopped) at $-\phi_{\text{FB}}$ or just before $-\phi_{\text{FB}}$. When approaching the flatband potential, one has to be careful that the current does not rise too high and causes damage to the sample. The scan rate is usually chosen based on the integration time of the FRA and the required spacing between the data points. Typical values range between 1 and 20 mV/s. The longer the FRA integration time, the more reliable the data is and the less noise one sees in the measurement. For modulation frequencies of a few hundred Hz and more, an integration time between 0.5 and 1 s generally gives good results. The amplitude of the modulation signal generated in the FRA is often chosen to be relatively large (0.1–1 V), and is then attenuated by a factor of 10, 50, or 100 inside the potentiostat. This is done to minimize noise pick-up in the cable between the FRA’s generator output and potentiostat’s modulation input. The modulation amplitude imposed on the sample should be as low as possible to avoid nonlinearities. Since the current–voltage characteristics of semiconductors typically vary as $\text{Exp}(eV/kT)$, the amplitude should be much smaller than 25 mV.  

Amplitudes higher than 10 mV should be avoided, and a value as small as 5 mV often gives good results. To ensure that the potentiostat is fast enough, its internal filters – such as the capacitor across the range resistor in Fig. 3.6 – should be switched off. This is crucially important, and easy to forget when writing one’s own program to do these measurements. In some potentiostats the bandwidth of the power amplifier itself can be changed.

---

17 When writing one’s own measurement software, it should be realized that the FRA simply divides the voltage signals at both input channels – it does not know what these signals represent. Since the (voltage) signal that represents the current is given by $I_{\text{monitor}}/C_2$ (Fig. 3.6), the FRA reports the impedance as $V/(I \times R_i)$ – it does not “know” the value of the range resistor $R_i$. The software should therefore multiply the reported value by $R_i$ to obtain the actual impedance.

18 The value of $kT/e$ at room temperature is ~25 mV.
For impedance measurements, one should select the fastest bandwidth model. The performance of the FRA can be optimized by using ac coupling for the channel 1 and channel 2 inputs. This inserts a capacitor in the signal line and thereby subtracts the dc component of the signal and improves the FRA’s ability to analyze small signals in a large background. Note that this setting should only be used for frequencies that are at least $10 \times$ higher than the -3 dB point of the ac-coupled input. The -3 dB point is usually below 10 Hz and can be found in the specifications.

### 3.6.5 Recognizing Instrumental Limitations

The most critical parameters to choose are the FRA’s modulation frequency and the potentiostat’s current range. A more sensitive current range gives higher quality data (higher accuracy, less noise), but it also makes the potentiostat slower. This can be understood from the diagram in Fig. 3.6, where the range resistor $R_i$ is seen to limit the speed at which the sample’s capacitance can be charged – essentially forming a low-pass filter. To ensure that the measured impedance does not depend on the potentiostat’s settings (the data should only reflect the properties of the sample), the potentiostat needs to be fast enough to follow the modulation. The best way to verify this is to measure a quick series of impedance spectra at different values of the potentiostat’s current range. As explained below, these measurements should be carried out at potentials not too far from the flatband potential. Fig. 3.16a shows simulated impedance spectra for the $R(RC)$ circuit shown in Fig. 3.12c for three values of the range resistor, whose values correspond to current ranges of 10 mA, 1 mA, and 100 μA. For the 100 μA current range ($R_i = 10 \ \text{kΩ}$), the imaginary part of the impedance drops steeply at high frequencies due to the limited bandwidth of the potentiostat. This can be avoided by choosing a higher current range, but this also results in smaller signals and higher noise levels. If the signals are not too small this can be remedied by choosing a longer integration time. Longer times than ~2 s are, however, not practical as they result in very long measurements or very widely spaced data points.

The effect of bandwidth limitations on a Mott–Schottky measurement is illustrated in Fig. 3.17a. The data shown in this figure are simulated for a current range of 100 μA ($R_i = 10 \ \text{kΩ}$). The other parameters are the same as those used to simulate the 10 kΩ curve of Fig. 3.16a. At 100 Hz and 10 kHz, the latter figure shows a perfectly linear relation between Log($-Z_{lm}$) and Log($f$), with a slope of $-1$ that corresponds to an ideal capacitive response ($-Z_{lm} = (\omega C)^{-1}$). The resulting Mott–Schottky plots for these frequencies are therefore identical. At 50 kHz, however, the Mott–Schottky plot is significantly different, and at 100 kHz the data are entirely unreliable. This shows that one should stay well away from the frequency region at which the slope of Log($-Z_{lm}$) vs. Log($f$) starts to deviate from $-1$. In addition to the value of the range resistor $R_i$, other parameters also affect the effective bandwidth of the system. As shown in Fig. 3.16b, the bandwidth improves when the width of the depletion layer ($W$) increases. This is because a larger value...
Fig. 3.16  Simulated impedance spectra (Bode plots) for the $R(\text{RC})$ circuit from Fig. 3.12c. The simulation includes a simple model that describes the bandwidth limitation of a (hypothetical) potentiostat. Several parameters are varied: (a) the range resistor, $R_i$; (b) the depletion layer thickness $W$; (c) the sample diameter $D$; (d) the resistance parallel to the space charge capacitance, $R_{\text{SC}}$. The series resistance $R_S$ is fixed at 140 $\Omega$ and $C_{\text{SC}}$ is calculated using $C_{\text{SC}} = (\varepsilon_0 \varepsilon_r A/W)$, where $\varepsilon_r$ is the dielectric constant (25) and the surface area $A$ is given by $1/4\pi D^2$. Note that the simulation in (a) does not show the increase in noise that one expects at less sensitive current ranges.

Fig. 3.17  Simulated Mott–Schottky plots for an n-type semiconductor. The same model is used as in Fig. 3.16, except that the depletion layer thickness is now calculated from (2.41) using a donor density of $10^{18}$ cm$^{-3}$. Left: Calculated curves using parameters that correspond to the $R_i = 10$ k$\Omega$ curve in Fig. 3.16a, showing the deviations that can occur at high frequencies when a (too) low current range limits the bandwidth of the potentiostat. Right: Deviations at low frequencies due to the presence of $R_{\text{SC}}$ parallel to $C_{\text{SC}}$, based on the parameters used in Fig. 3.16d.
of $W$ gives a smaller space charge capacitance, and a concomitantly smaller time constant for the low pass filter formed by $R_i$ and $C_{\text{sample}}$. This also implies that one should check the bandwidth limitations of the system by measuring the impedance spectra at potentials close to the flatband potential, i.e., the potential regime where $C_{\text{SC}}$ has its highest value.

A very effective method to increase the available frequency window is by using smaller samples. The effect of the effective surface area is shown in the simulation of Fig. 3.16c. A reduction in sample diameter by a factor of two decreases the capacitance by a factor of 4, i.e., it scales quadratically.

### 3.6.6 The Reference and Quasi-Reference Electrodes

In addition to the potentiostat and the sample, the reference electrode may also affect the effective bandwidth of the system. This is because the impedance of the reference electrode forms a low pass filter with the input capacitance of the potentiostat’s electrometer. For example, a 100 kΩ reference electrode combined with a 5 pF input capacitance (cf. Sect. 3.3.3) gives −3 dB point of ~0.3 MHz ($f = (2\pi RC)^{-1}$), which would limit the frequency for accurate measurements to ~30 kHz. The impedance of the reference electrode mostly depends on the type of diaphragm or membrane that is used. A highly permeable membrane gives a low impedance, but also risks contamination of the reference electrode’s internal solution with ions from the electrolyte of the PEC cell. It should also be realized that the impedance may increase over time when the membrane get clogged.

One can measure the impedance of a reference electrode using impedance spectroscopy. This can be done using a simple two-electrode configuration, in which the reference electrode is connected as the working electrode (!), and the counter electrode is a regular Pt counter electrode. The electrolyte should be highly conducting, e.g., a 1 M KOH or NaCl solution. One should be careful to keep the applied bias potential at 0 V at all times to avoid large dc current flows through the reference electrode – this may destroy it. The impedance spectrum is measured with a 5 or 10 mV modulation amplitude, over a frequency range of, e.g., 100 Hz to 10 kHz. The imaginary part of the impedance should be very small, and the real part should be constant over a fairly large range. If this is indeed the case, the real part of the impedance is the impedance of the reference electrode.

For high-impedance reference electrodes, or for potentiostats with a high input capacitance, one should consider using a so-called quasi-reference electrode. This is a piece of Pt wire that is immersed in the electrolyte next to the normal reference electrode, and connected to the reference electrode via a small capacitor, as shown in Fig. 3.15. The capacitor blocks any dc signals, so for constant or slowly varying

---

19 This membrane forms the junction between the internal solution of the reference electrode and the electrolyte of the PEC cell.
potentials the potentiostat only “sees” the normal reference electrode. At high frequencies, however, the impedance of the capacitor decreases and the potentiostat only senses the potential via the Pt wire. The impedance of the capacitor is given by \( Z = (2\pi f C)^{-1} \), and \( C \) should be dimensioned so that the impedance is less than 1–10 k\( \Omega \) at the frequencies of interest. A value of 100 nF gives good results in almost all cases. Using a quasi-reference electrode for impedance measurements is always a good idea, even if the reference electrode has an acceptably low impedance. It has no undesired side-effects, and it ensures that the high-frequency data remains reliable if the resistance of the reference electrode membrane increases over time (due to, e.g., clogging).

### 3.6.7 Sample Limitations and Complications

Even if there are no instrument-related limitations, obtaining a representative Mott–Schottky plot can still be a major challenge – or even impossible – due to the nature of the sample itself. For example, if the equivalent circuit of Fig. 3.12b is too simple to model the sample’s behavior, additional circuit elements are necessary which introduce cross-terms in the expressions for the real and imaginary parts of the impedance. The space charge capacitance can then no longer be obtained by measuring \( Z_{\text{Im}} \) (and/or \( Z_{\text{Re}} \)) at a single frequency. Instead, one has to measure a series of impedance spectra at various applied potentials, and then identify a single equivalent circuit that adequately describes all these spectra. The values of all the elements in this circuit can then be determined as a function of the applied potential from a nonlinear least-squares fit of each individual impedance spectrum. The final step is to construct a Mott–Schottky plot from the fitted values of \( C_{\text{SC}} \) vs. the potential. Clearly, this is a complicated procedure that one usually tries to avoid if the goal is simply to measure the donor density or the flatband potential. There are several possible reasons why a simple \( RC \) series circuit may not be sufficient to describe the sample’s behavior. A few of the most important ones are described below.

**Parallel resistance \( R_{\text{SC}} \)** – If the semiconductor film contains a large number of defects, or has pin-holes that expose the underlying back-contact to the electrolyte, this is often accompanied by a relatively large steady-state background current. This can be modeled by a resistance \( R_{\text{SC}} \) in parallel with the space charge capacitance, as illustrated in Fig. 3.12c. The presence of \( R_{\text{SC}} \) mainly affects the real part of the impedance at low frequencies. While it does not seem to influence the imaginary part of the impedance all that much, the simulation in Fig. 3.17b shows that significant deviations of the Mott–Schottky curve can still occur at low frequencies for sufficiently small values of \( R_{\text{SC}} \). Deviations of this kind can be avoided by avoiding low frequencies and by improving the quality of the sample.

**Surface States** – The presence of surface states introduces one or more additional \( RC \) series elements in the equivalent circuit, as illustrated in Fig. 3.12a. If the
energy level of the surface states is far away from the band edges, surface state charging will be slow and ignored by choosing sufficiently high modulation frequencies. While the physical or chemical origin of surface states is often difficult to determine and poorly understood, their contribution to the impedance of the photoelectrode is well established [52–54]. Specifically, the presence of surface states is known to cause frequency dispersion in the Mott–Schottky plot. When considering only the space charge and surface state capacitances and resistances, the influence of the surface state capacitance on the overall capacitance of the sample can be estimated with [53]:

\[
C_{\text{eff}} = C_{\text{SC}} + \frac{C_{\text{SS}}}{1 + (\omega R_{\text{SS}} C_{\text{SS}})^2}.
\]  

(3.11)

The density of surface states at the Fermi level, \(N_{\text{SS}}(E_F)\), can be determined by measuring an impedance spectrum and extracting the value of \(C_{\text{SS}}\), which is given by \(C_{\text{SS}} = eAN_{\text{SS}}(E_F)\) [53]. By doing this for a range of potentials, one can determine the energetic distribution of surface states over the energy gap. This may help to design passivation strategies for surface states in order to suppress recombination. Depending on their nature and energetic position, surface states can sometimes be passivated by suitable surface treatments, e.g., by depositing a few atomic layers of \(\text{Al}_2\text{O}_3\) or \(\text{SiO}_2\) on top of the photoelectrode’s surface.

**Helmholtz capacitance** – For samples with very high donor densities, the space charge capacitance may become so large that it is no longer possible to ignore the Helmholtz capacitance, \(C_H\), which has a constant value of 10–20 \(\mu\text{F/cm}^2\). The effect on the Mott–Schottky curve is a shift of its slope to more negative potentials (for photoanodes). This means that the slope itself is unaffected and can still be used to determine the donor density. The flatband, however, shifts by an amount given by [52]

\[
\Delta \phi_{FB} = -\frac{\varepsilon_0 \varepsilon_r e N_D}{2C_H^2},
\]  

(3.12)

It should be noted that extremely high donor densities (>\(10^{20}\) cm\(^{-3}\)) are required for an appreciably shift of the (apparent) flatband potential.

**Low carrier concentrations** – In some cases, the background (or bulk) concentration of free carriers is so low that the entire film is depleted at all values of the applied potential. This means that the semiconductor film effectively behaves as an insulator with a constant (potential-independent) capacitance. This is not often observed for metal oxide photoelectrodes, which usually have free carrier concentrations between \(10^{17}\) and \(10^{19}\) cm\(^{-3}\) due to either impurities, intrinsic defects, or intentionally introduced dopants. However, there are exceptions. Unpublished results from the author’s laboratory show that \(\text{TiO}_2\) films deposited by spray deposition of high-purity TTIP (99.99%) dissolved in ethanol yield insulating films. A different situation
occurs if the donor or acceptor level is relatively far away from the band edge, and ionization of these species is slow.\textsuperscript{20} Under these conditions, the Mott–Schottky plot shows a horizontal curve for high frequencies, similar to what is observed for an insulator. At low frequencies, however, the carriers can follow the modulation voltage and a normal Mott–Schottky curve is found. This situation has been observed for spray-deposited InVO\textsubscript{4} photoanodes [55].

Nanostructured electrodes – a final important example is that of nanostructured porous electrodes. Such mesoporous morphologies are used to increase the active surface area (better catalysis), improve light absorption through scattering, and/or to provide short diffusion paths for the photogenerated minority carriers to the electrolyte. The latter is essential in the case of, e.g., \(\alpha\)-Fe\textsubscript{2}O\textsubscript{3} due to its very small minority carrier diffusion length (see Chap. 4). The impedance of mesoporous semiconductor photoelectrodes can be described by the so-called transmission lines, which are basically ladder networks of resistive and capacitive elements. The analysis of such systems is a fairly complicated topic. Much progress has recently been made in this area, most notably by Bisquert et al., and for a more in-depth discussion the reader is referred to the literature [47, 56].

Under certain conditions it may still be possible to determine the donor density of a porous photoelectrode from a Mott–Schottky measurement. This is the case for high donor densities, when the space charge width is small and is therefore still able to track the surface contours. An example of this for nanostructured Si-doped \(\alpha\)-Fe\textsubscript{2}O\textsubscript{3} photoanodes is shown in Fig. 3.18, reported by Cesar et al. [57]. The actual surface area was estimated by dye absorption experiments, and the estimated donor density for this system was \(\sim10^{20}\) cm\textsuperscript{-3}. The concave shape of the Mott–Schottky plot is consistent with a gradual decrease in the effective surface area (cf. (3.5)) as the potential is increased and the depletion layer progressively penetrates into the bulk [58].

3.6.8 Nonideal Space Charge Capacitance: The Constant Phase Element

Even if the depletion layer contribution can be successfully separated from the other contributions in the system, one may find that it does not behave as an ideal capacitor. A clear indication of this is the observation of frequency dispersion in the Mott–Schottky plot, or a somewhat depressed semicircle instead of a perfectly round one in a Nyquist plot of the impedance spectrum.\textsuperscript{21} In such cases, it is often

\textsuperscript{20}The donors/acceptors are then no longer shallow, but deep.

\textsuperscript{21}Note that a semicircle in the \(Z'\) vs. \(Z''\) (Nyquist) plot is only observed in the presence of a resistive element in parallel to the space charge capacitance.
possible to describe the space charge response by a so-called Constant Phase Element (CPE). The impedance of a CPE is given by\(^{22}\)

\[
Z_Q = Y_Q^{-1} (j\omega)^{-n} \quad \text{(3.13)}
\]

Depending on the value of \(n\), the electrical behavior of a CPE can range between purely resistive (\(n = 0\)) and purely capacitive (\(n = 1\)). \(Y_Q\) is often interpreted as the numerical value of the capacitance of the element. This is only true for \(n = 1\), since for lower values of \(n\) the units are (\(\Omega^{-1} \text{s}^n\)) instead of Farads (\(\Omega^{-1} \text{s}\)). This means one cannot directly calculate the characteristic time constant, \(\tau\), for an \(RQ\) circuit as one would normally do for an \(RC\) circuit (\(\tau = RC\)). Hsu and Mansfeld showed that the equivalent capacitance of a CPE can be written as \(^{59}\):

\[
C_Q = Y_Q (\omega_{\text{max}})^{n-1} \quad \text{(3.14)}
\]

Here, \(\omega_{\text{max}}\) is the frequency at which the imaginary value of the impedance has a maximum, and \(C_Q\) has the proper dimensions of Farads (\(\Omega^{-1} \text{s}\)). It should be noted

\(^{22}\) CPE elements are usually indicated by the letter \(Q\), analogous to the letters \(R\) and \(C\) used for resistive and capacitive elements, respectively.

---

**Fig. 3.18** Mott–Schottky plots of Si-doped (curves \(a–c\)) and undoped (curve \(d\)) mesoporous hematite photoanode. The capacitances for curves \(a\), \(b\), and \(c\) are obtained from an Si-doped sample and models \(a\), \(b\), and \(c\), respectively, shown in the inset of the left-hand plot. Curve \(d\) is obtained from the undoped film and model \(a\) (series \(RC\)). The dashed lines connecting the data points represent the variable active surface area fit. Sketches \(e–g\) depict the development of the space-charge layer in a mesoporous semiconductor as function of applied potential, illustrating a decrease in active surface area at advancing space-charge layer width in two dimensions. \((e)\) Near flat band potential with maximum surface area. \((f)\) Total depletion of smaller feature at increased bias potential. \((g)\) Decreased active surface area in concave curved surface. Reprinted with permission from ref. [57], copyright, 2009 American Chemical Society.
that \( \omega_{\text{max}} \) can only be determined if the Nyquist plot shows a semicircle, i.e., if there is a resistive element present in parallel to the CPE element.

The physical origin of nonideal, CPE-like behavior of the space charge capacitance is often unclear and difficult to determine. A CPE element generally describes processes that have a distribution of time constants or activation energies, so it could describe a collection of surface or bulk states with energy levels that are distributed over the bandgap.\(^{23}\) Dielectric relaxation phenomena in a thin surface layer of the semiconductor or at the semiconductor/electrolyte double layer have also been proposed as possible causes for frequency dispersion \([60, 61]\). In the case of porous electrodes (Sect. 3.6.7), the capacitive element of the transmission line model can be modeled as a CPE element \([47]\). Finally, current inhomogeneities due to surface roughness effects are often mentioned as an explanation for CPE-like behavior. For such electrodes, however, the frequency dispersion has been reported to disappear at the high electrolyte concentrations (\(>0.5\) M) that are commonly used in PEC research \([61]\).

Acknowledgments  The author gratefully acknowledges the NWO-ACTS Sustainable Hydrogen program (project 053.61.009) and the European Commission’s Framework 7 program (NanoPEC, Project 227179) for support.

References


\(^{23}\) A bulk or surface state with a single, discrete energy level is probably better described by a single additional \(RC\) component instead of a CPE element.


Part II

Materials Properties and Synthesis
Chapter 4
Nanostructured $\alpha$-$\text{Fe}_2\text{O}_3$ Photoanodes

Kevin Sivula

4.1 Introduction

Since the seminal demonstration of photocatalytic hydrogen production from water splitting with TiO$_2$ [1], scientists and engineers have sustained the search for a photoelectrochemical (PEC) water-splitting material combining efficient solar light harvesting, high quantum efficiency, practical durability, and low cost. However to date, no single semiconducting material which meets all of these requirements has been found [2]. While systems delivering high solar-to-hydrogen (STH) efficiency have been demonstrated using III–V semiconductor materials with optimized band gaps and energy levels straddling the hydrogen and oxygen redox potentials their cost and stability are a major disadvantage [3]. In contrast, many transition metal oxide semiconductors offer suitable stability to the chemically harsh environment of water, oxygen, and light inherent to PEC water splitting. For example, TiO$_2$ and WO$_3$ have been thoroughly investigated and optimized to deliver high single-wavelength quantum efficiencies [4, 5]. Unfortunately the stability of these oxides in aqueous environments, aided by an energetically low valence band which is defined largely by O 2p energy levels, also brings a large bandgap which renders these materials transparent to most of the incident solar illumination. This limits the maximum possible STH conversion efficiency. Specifically, considering the bandgap energies, $E_g$, alone, TiO$_2$ (anatase, $E_g = 3.2$ eV) and WO$_3$ ($E_g = 2.6$ eV) could only convert 3.4 and 10.2%, respectively, of the sun’s energy into hydrogen. A more accurate evaluation of the potential efficiencies, including expected kinetic losses, suggests that only 5% STH is possible with WO$_3$ [6].

In addition to sufficient light absorption and stability, a semiconducting material used for PEC water splitting must be made from raw materials with ample availability.
to be relevant for energy conversion on a global scale. Iron(III) oxide is a promising material in light of these requirements. With a potential to convert 16.8% of the sun’s energy into hydrogen [6], it has been extensively examined for application to solar water splitting. In this chapter the advantageous properties of Fe₂O₃ as well as the challenges it presents are examined in the context of PEC water splitting. The most recent efforts at improving photon harvesting by nanostructuring are reviewed, and continuing efforts to control the surface catalysis, and enhance the electronic properties by doping, are also addressed by examining recent literature reports in the field.

4.2 Iron(III) Oxide: A Promising Material for Solar Hydrogen Production

As iron is a readily oxidized metal, and iron ions can exist in different charge states (mainly 2+ and 3+) separated by relatively small energy differences, there are many forms of iron oxide both found in nature and that can be synthesized easily in the lab. In addition, because iron is a common element in the earth’s crust (6.3% by weight) iron oxide is everywhere around us, giving, for example, the geological formations of the southwest USA their brown-red color. Commonly (and perhaps unfortunately), iron oxide is better known as a nuisance material, and by its common name – rust – the red-brown corrosion product of iron in the presence of oxygen and moisture. However, its chromatic characteristics also imply the ability to absorb solar irradiation. This coupled with its abundance and nontoxicity make iron oxide a particularly attractive material for use in solar energy conversion. The distinct properties of all of the iron oxides are comprehensively presented in a recent book by Cornell and Schwertmann [7]. Here, a brief summary is presented of the important and unique properties of iron(III) oxide, α-Fe₂O₃, or hematite, the most frequent iron oxide employed for solar hydrogen production. These properties highlight both the promising aspects and the challenges that hematite presents as a material for solar energy conversion.

4.2.1 Structural and Material Properties

For iron oxide in its fully oxidized form, each iron atom is in the +3 oxidation state. There are thus two O²⁻ anions for every three Fe³⁺. Even considering this stoichiometry, there are many types of Fe₂O₃ including amorphous structures. The α (hematite) and γ (maghemite) forms are two most common crystalline structures found in nature. Maghemite can be considered to be a form of magnetite (Fe₃O₄), a ferrimagnetic spinel iron(II–III) oxide, with cation vacancies. Indeed, natural maghemite arrives from the oxidation of magnetite, and retains strong (ferri) magnetic properties. While this form of iron oxide has found wide-spread use as
a magnetic recording medium (giving floppy disks and cassette tapes their brown color) the physical properties have not been found to be suitable for photoelectrochemistry. In addition, maghemite, as with all of the other iron oxides converts to hematite, the most thermodynamically stable form of iron oxide (under standard conditions) upon treatment with heat.

Hematite, alpha-Fe$_2$O$_3$, is the most common form of iron oxide. Due to its prevalence, hardness (6.5 on Mohs scale), and chemical and thermal stability (decomposition temperature of 1,566°C), it has found historic use as a polishing powder known as jeweler’s rouge. In addition, its ability to absorb light has made it an attractive choice for pigments in paints and even cosmetics. Alpha-Fe$_2$O$_3$ has the same crystal structure of corundum, Al$_2$O$_3$, which is trigonal (hexagonal scalenohedral, symbol $\overline{3}2/m$) with space group $R\overline{3}c$ with lattice parameters $a = 5.0356$ Å, $c = 13.7489$ Å, and six formula units per unit cell [7]. It is easy to understand hematite’s structure based on the packing of the anions, O$_2^-$, which are arranged in a hexagonal closed-packed lattice along the [001] direction. The cations, Fe$^{3+}$, occupy the two-thirds of the octahedral interstices (regularly, with two filled followed by one vacant) in the (001) basal planes. The tetrahedral sites remain unoccupied. Two views of the hematite crystal structure are shown in Fig. 4.1. Here the layered structure of iron oxide is evident; layers of closed-packed oxygen anions alternate with iron cations bilayers in the [001] direction. The arrangement of cations can also be thought of producing pairs of FeO$_6^{9-}$ octahedra which share edges with three neighboring octahedra in the same plane and one face with an octahedron in an adjacent plane in the [001] direction. This face sharing is responsible for a trigonal distortion of the octahedron. In the FeO$_6^{9-}$ octahedra which share faces, iron atoms are repelled along the direction normal to the [001] axis. Consequently, the O–O distance in shared faces is shorter than in an unshared face. However, the electronic structures of these distorted FeO$_6^{9-}$ octahedral are quite similar to undistorted clusters [7].

![Fig. 4.1](image-url) The crystal structure of hematite as seen looking down the [001] direction (left) and the [8] direction (right). Oxygen anions are dark gray and iron cations are light gray. The unit cell is outlined, and the orientation map is shown for each view.
4.2.2 Magnetic Properties

The arrangement of the oxygen anions and the high-spin (d5 configuration) iron cations in iron(III) oxide naturally affects the orientation of the iron atoms’ spin magnetic moment and thus the observed bulk magnetic properties. The diverse magnetic properties of all the iron(III) oxides have been recently reviewed [9]. In brief, the inverse spinel structure of maghemite (γ-Fe2O3) results in the presence of two interpenetrating sublattices and gives rise to ferrimagnetic behavior and high susceptibilities. In contrast, hematite is antiferromagnetic (temperatures below 260 K) or a weak (parasitic) ferromagnet at room temperature. The latter is due to the ferromagnetic coupling of the spins within the (001) basal (iron bilayer) planes and antiferromagnetic coupling between iron bilayers along the [001] direction [10]. Here the trigonal distortion of the FeO6 octahedra causes the spins to become slightly canted (about 5°) causing the destabilization of their perfectly antiparallel arrangement (parasitic ferromagnetism). Above the Néel temperature (950 K), the iron spins become thermally disordered and a paramagnetic material results [9]. While the magnetic properties of hematite are not particularly pertinent to its photoelectrochemical performance, the iron spin configuration does influence the optoelectronic and carrier transport properties of hematite. In addition, exploiting magnetic characteristics with Mössbauer spectroscopy to identify phases in iron oxides prepared for electrochemical application has proven to be a useful technique [11–13].

4.2.3 Optoelectronic Characteristics

The dense (hexagonal close) packing of oxygen combined with the interstitially positioned iron yields a very dense structure (5.26 g/cm³) exhibiting a high polarizability and a high refractive index (3.15) [14]. This layered structure also generates complex behavior when interacting with photons and electrons. These attributes of hematite should be properly understood for application as a semiconductor for solar water splitting. Decades of research have passed scrutinizing the mechanisms of charge transport and light absorption, and while hematite is still not completely understood, much progress has been made.

The absorption of photons by hematite is shown in Fig. 4.2 (left side) and begins in the near infrared spectral region. This initial weak absorption is due to d–d transition states between electron energy levels of the d5 Fe3+ ion, which are split by an intrinsic crystal field [16]. Two absorption bands in this spectral region centered around λ = 860 and 640 nm with absorption coefficients, α, of the order 10³ cm⁻¹ have been accurately assigned using ligand field theory to spin-forbidden transitions from 6A1g to 4T1g and 4T2g, respectively [15, 16]. While photoexcitation of hematite at these wavelengths has been shown to increase its conductivity [15], sustained photocurrent is not observed in a PEC system (see Sect. 4.2.4) upon
irradiation below the bandgap energy, \( E_g \) (which, depending on the method of preparation of hematite, is usually reported to be between 1.9 and 2.2 eV or 650–560 nm [14]). The absorption coefficient of pure hematite increases abruptly at the band gap energy and further up the electromagnetic spectrum \( \alpha \) continues to increase with additional absorption peaks having been observed centered on 2.4, 3.2, and 5.8 eV (with \( \alpha \) on the order of \( 10^5 \) cm\(^{-1} \)) in samples of polycrystalline hematite [17]. Hematite’s strong absorption of yellow to ultraviolet photons in the visible region and transmission of orange to infrared photons gives it a characteristic red color. However, it is fairly uniform reflectivity as a function of visible light wavelength gives polished bulk samples of hematite a metallic look.

Since the electronic nature of the band gap in hematite is of great interest to understand its performance as a material for solar energy conversion, much work has focused on this aspect. Tauc analysis of the bandgap absorption onset, which assumes that the energy state bands are parabolic with respect to the crystal momentum, most frequently indicates an indirect (phonon-assisted) band gap transition [18]. However, a few recent reports of a direct band gap in hematite have been attributed to quantum size-effects [19, 20]. The initial orbital

Fig. 4.2 Optical and electronic properties of hematite: The left graph shows the absorption coefficient as a function of photon wavelength in hematite thin films prepared by r.f. sputtering (from [15] used with permission). The right graph shows the calculated density of electron states (spin up and down) in hematite. The Fermi energy is at 0 eV and two lines at about −5.5 and 1.8 eV divide the plot into three regions: the lowest energy region contains the occupied Fe-d states; the next energy region contains the occupied O-p states; the region above the gap contains the empty Fe-d levels (Reprinted figure with permission from Velev, J., Bandyopadhyay, A., Butler, W.H., Sarker, S.: Phys. Rev. B: Condens. Matter 71, 205–208 (2005). Copyright 2005 by the American Physical Society)
assignments of the bandgap suggested it was due to an indirect transition of Fe$^{3+}$ d–d origin [15, 21], and that a stronger direct transition involving a charge transfer from an O 2p orbital to Fe 3d did not occur until 3.2 eV [21]. This led to a hypothesis that two different types of p-type charge carriers (holes) could be produced in hematite, depending on the excitation mechanism, and were responsive for the observed difference in PEC performance as a function of wavelength [22, 23]. However, the most recent ab initio calculations to determine the electronic structure of hematite by the Hartree–Fock approach [24] and density functional theory with a local spin-density approximation and coulomb correlation [25, 26] both predict that the highest occupied energy states are primarily O p in character and the lowest unoccupied states are from an empty Fe-d band (see Fig. 4.2, right). This conclusion is also supported by soft-X-ray (O K-edge) absorption and emission spectroscopy, which when compared to photoemission spectra from configuration-interaction FeO$_6$ cluster calculations, confirm that the valence band is at least strongly hybridized and indicates further that it is mostly of O 2p character [27]. These latest results not only suggest that pure stoichiometric hematite is a charge transfer insulator, and not a Mott–Hubbard type insulator – a detail that affects conduction models – but also contradicts the suggestion that two types of holes, originating from different transitions, could cause different PEC behavior.

Finally, as discussed in the next section, much effort has been placed on doping hematite with various elements to increase charge carrier conduction. These elements are typically present at 1 atom% or less, and while substitutional transition metal impurities have been predicted to introduce interband gap energy states [25], the same dopants have not been reported to significantly change the bandgap absorption energy or the absorption coefficient [28–31].

### 4.2.4 Charge Carrier Transport and Dynamics

During an electromagnetic energy absorption event in hematite, a photon with energy greater than the bandgap will excite an electron from an occupied to an unoccupied state. For this process to be useful for water splitting the excited electron and remaining electron vacancy (hole) must spatially separate and live long enough to reach points of their collection or reaction. Luminescence studies on nanoparticles show that the excitation of hematite with photons of 3.2 eV results in fluorescence quantum yields of about $1 \times 10^{-5}$ [32]. This extremely low value of radiative carrier recombination suggests that fast nonradiative processes such as carrier trapping and phonon coupling are limiting the excited state lifetime [33]. Indeed, the ultra fast dynamics of excited states in hematite have been studied in nanoparticles, thin films, and single crystals using femtosecond laser spectroscopy [32, 33]. In nanoparticles it was found that the excited state decay profiles were independent of the pump power and the probe wavelength, and were not affected by lattice doping or surface modification. However, the lifetime of the excited state was determined to be very short – 70% of the transient absorption had disappeared
after just 8 ps and could not be detected after 100 ps [32]. Studies on epitaxially grown thin films (100 nm using an oxygen plasma-assisted molecular beam) reduced the influence of defect states in the bulk or on the surface compared to the nanoparticle case, but found similar excited state dynamics, which were described by the following scenario: Initially, hot electrons relax to the conduction band edge within 300 fs. Then their recombination with holes and trap states occurs within 3 ps. The resulting trap states can exist for hundreds of picoseconds or longer [33]. In addition, the dominant carrier trapping mechanism was ascribed to mid-gap Fe$^{3+}$ d–d states. These carrier trap states have been reported to be 0.5–0.7 eV below the conduction band edge giving them an optical transition of about 1.5 eV [34, 35] and are thus likely the same spin-forbidden transitions intrinsic to hematite which were discussed in the previous section.

A common metric sought by researchers in the field of semiconductor solar energy conversion is the excited carrier diffusion length which describes how far an average excited state can travel before recombination, and is estimated by taking the square root of the product of the carrier lifetime and the diffusion coefficient. Thus an ultra short lifetime alone does not prevent high efficiency in a semiconductor for PEC energy conversion if the carriers have a high mobility through the crystal lattice. Accordingly, understanding the transport of carriers in hematite is equally important to understanding the lifetime.

The earliest work by Morin on pure sintered polycrystalline hematite describes very low electrical conductivities (ca. $10^{-14}$ $\Omega^{-1}$ cm$^{-1}$) [36], conduction electron concentrations of $10^{18}$ cm$^{-3}$ at 1,000 K, and an activation-energy type electron mobility on the order of $10^{-2}$ cm$^2$/V/s [37]. Studies on pure single crystals showed conductivities below $10^{-6}$ $\Omega^{-1}$ cm$^{-1}$ and were not suitable for further electronic characterization [31]. These unusually small values obliged electrical conduction to be explained by Fe$^{3+}$/Fe$^{2+}$ valence alternation on spatially localized 3d orbitals. Modeling studies have had success matching empirical data by describing the conduction mechanism with a small polaron model which includes the effect of the larger size of the Fe$^{2+}$ ion and the associated lattice distortion (polaron) [38, 39]. Conduction of electrons or holes can then be best described by hopping of polarons with an activation energy. This conduction mechanism causes the mobility of carriers to increase with increased temperature, as transport is assisted by lattice vibrations [40]. The localized electron, small polaron model has also been further supported by experimental results including pulse radiolysis [41] and the previously mentioned femtosecond laser spectroscopy on nanoparticles [32].

Further studies of hematite single crystals identified a highly anisotropic electron transport with conduction within the iron bilayer (001) basal plane four orders of magnitude greater than perpendicular directions (parallel to [001]) [42, 43]. This discrepancy cannot be explained by the proximity of iron cations alone, as the shortest Fe–Fe distances are actually along the [001] direction and nevertheless they are quite similar to Fe–Fe distances in the basal planes [44]. However, the anisotropic conductivity can be classically explained considering Hund’s rule and the magnetic structure of hematite. The ferromagnetic coupling of the spins in the (001) basal planes and antiferromagnetic coupling along the [001] direction create
an environment where electrons can move (n-type conductivity) within the iron bilayers (an environment of parallel spins) but are forbidden to hop across the oxygen plans to an iron bilayer with the opposite spins. Thus conduction in the [001] direction could only involve the movement of holes in the form of Fe\(^{3+}\) → Fe\(^{4+}\) electron transfer (p-type conductivity). This process is significantly slower in hematite [40]. While this classical explanation is generally accepted, recent ab initio electronic structure calculations combined with electron transfer (Marcus) theory have correctly predicted the large transport anisotropy in hematite but suggested it arises from the slowness of both hole and electron mobility across basal oxygen planes [10]. This is contrast to the classical conception as it does not consider electron transport to be forbidden in the [001] direction, but instead identifies the most important factor that influences the carrier mobility to be the electronic coupling – a quantity found to depend on both a superexchange interaction between the bridging oxygen atoms and the d-shell electron spin coupling.

Regardless of the specific details of the conduction mechanism it is clear that the conduction properties of pure hematite are not appropriate for use as a PEC semiconducting material. However, the conductive properties of hematite can be significantly enhanced by adding impurities to act as electronic dopants. Indeed, it is possible to increase conductivities and obtain both p-type and n-type alpha-Fe\(_2\)O\(_3\) by doping with atoms such as (but not limited to) Mg\(^{2+}\), Cu\(^{2+}\) (p-type) or with Ti\(^{4+}\), Sn\(^{4+}\), Zr\(^{4+}\), Nb\(^{5+}\) (n-type) [40]. This behavior can be easily explained if we consider hematite to be a small polaron conductor with n- and p-type charge carriers assumed to be Fe\(^{2+}\) and Fe\(^{4+}\), respectively. Then, if the impurity atoms substitute for a Fe\(^{3+}\) ion in the crystal lattice, they would cause the creation of Fe\(^{2+}\) (n-type) or Fe\(^{4+}\) (p-type) to balance the lattice charge. By doping at sufficient levels, high carrier conductivities can be attained. For example, Zr\(^{4+}\) was doped into single crystals to give donor densities on the order of 10\(^{19}\) cm\(^{-3}\), conductivities around 0.1 Ω\(^{-1}\) cm\(^{-1}\), and increased electron mobility (perhaps due to a increase in dielectric constant) of 0.1 cm\(^2\)/V/s [31]. A review by Shinar and Kennedy outlines the dopants used for n-type iron oxide and their affect on electronic and PEC properties [45].

In addition to foreign impurities acting as electronic dopants, oxygen vacancies play an important role in the electronic behavior of hematite, and are the most commonly present intrinsic defects found in hematite [17]. In fact, samples of natural specular hematite have been found to be degenerate semiconductors due to oxygen deficiencies [34]. This is because oxygen vacancies act as n-type dopants by creating Fe\(^{2+}\) ions. Elevated temperature treatment (above 800°C) in air or anaerobic environments has been shown to create oxygen vacancies in hematite [17]. In this case, the oxygen vacancies may result in the formation of a Fe\(_2\)O\(_4\) on the surface, distorting conduction measurements [38], and causing extensive recombination of carriers and reduced photoconduction that can be regained by oxygen treatment [46].

Finally, based on the lifetime and charge carrier mobility found in hematite, the diffusion length of excited carriers is expected to be small. Kennedy and Frese calculated the diffusion length for holes in n-type (Ti\(^{4+}\) doped) polycrystalline samples using the experimental photocurrent efficiency and a band-bending
semiconductor liquid interface model to determine the mean diffusion length for holes, \( L_p \), outside the space charge region as 2–4 nm \[18\], and a similarly indirect method gave the same order of magnitude for pure \( \text{Fe}_2\text{O}_3 \) [47]. It is important to note that this value is significantly smaller compared to other oxides used for PEC water splitting. For example, values of \( L_p \) up to 10⁴ and 150 nm have been reported for \( \text{TiO}_2 \) [48] and \( \text{WO}_3 \) [49], respectively. The ramifications of such a low value for \( L_p \) on the performance of hematite for PEC water splitting are discussed in the next section.

### 4.2.5 Photoelectrochemical Properties: Challenges for Solar Water Oxidation

Hardee and Bard [50] first turned to \( \text{Fe}_2\text{O}_3 \) as a material for water photolysis in 1976 seeking a photoanode material that was both stable under anodic polarization and capable of absorbing light with wavelengths longer than 400 nm. Before this, materials like GaAs, CdS, and Si had been found to absorb an appreciable amount of solar irradiation, but also to be unstable in aqueous environments under anodic condition. In contrast, materials like \( \text{TiO}_2 \), \( \text{SrTiO}_3 \), and \( \text{SnO}_2 \) had been identified as stable, but had bandgaps over 3 eV. Hardee and Bard made thin films of \( \text{Fe}_2\text{O}_3 \) on Ti and Pt substrates by chemical vapor deposition (CVD) of \( \text{Fe(AcAc)}_3 \) and observed photocurrent from their electrodes under illumination of 500 nm light at a potential of 0.8 V vs. SCE in 0.5 M KCl as electrolyte. Soon after Yeh and Hackerman [51] simply heated iron foils in air to make n-type iron oxide electrodes that gave photocurrents with no observable corrosion problems from pH 4.0 to 14 and Quinn et al. [52] showed that the (012) face of flux grown single crystals had collection efficiencies of around 20% at 0.5 V vs. SCE in 2 M NaOH (1.57 V vs. RHE) with 475 nm light. Further work by Kennedy et al. showed that p-type doping (with \( \text{Pb}^{2+} \) and \( \text{Ca}^{2+} \)) produced electrodes with high resistivity, but \( \text{Si}^{4+}, \text{Sn}^{4+}, \) and \( \text{Ge}^{4+} \) doped n-type hematite performed better than \( \text{Ti}^{4+} \)-doped hematite [28]. In addition, these initial studies showed that hematite had a flatband potential, \( V_{fb} \), too low in energy to reduce water, and an external bias had to be applied to complete the water-splitting reaction (the electron energy scheme for a n-type hematite photoanode water-splitting system is shown in Fig. 4.3, left). Despite this, hematite was still considered to be a promising n-type photoanode material for solar water splitting due to its capability to absorb a significant portion of solar irradiation (29% of AM 1.0 [52]) and superior stability. Over the next decade numerous reports were published describing continuing studies with pure and doped \( \text{Fe}_2\text{O}_3 \) made by various routes. Notable extensive PEC studies were performed on Ti-doped polycrystalline-sintered hematite pellets [18, 23], and Nb-doped single crystals [53] – the latter reporting champion performance with an incident photon conversion efficiency (IPCE) of about 37% at 370 nm and 1.23 V vs. RHE in 1 M NaOH (27% at 1.06 V, see Fig. 4.3, right).
A good review of these seminal studies on hematite has been given by Lindgren et al. [14] and in summary, the first decade of work identified the many challenges of employing this material for PEC water splitting:

1. Flatband potential too low in energy for water reduction [23, 52]
2. Large overpotential and slow water oxidation kinetics [23]
3. Low absorption coefficient, requiring relatively thick films [54, 55]
4. Poor majority carrier conductivity [35, 53]
5. Short diffusion length of minority carriers [18]

These limitations deterred much further interest in using hematite as a photoanode for water oxidation, as it appeared that the conversion of solar energy with high efficiency with it was not possible. Specifically, the small polaron transport and the deep inter-bandgap traps in hematite, which were discussed in Sect. 4.2.3, result in challenge numbers 4 and 5 above. Drawback 4 requires high doping levels to increase the ionized donor concentration and thus the conductivity. This in turn reduces the width of the space charge layer [56]. For example, donor dopant concentrations in the champion Nb-doped single crystals were found to be $5 \times 10^{19}$ cm$^{-3}$ by Mott–Schottky analysis [53]. Assuming classical depletion layer theory [56], this would result in a depletion layer width, $W$, of 7 nm at a band bending of 0.25 V and $\varepsilon = 100$. The low absorption coefficient in hematite (challenge 3) then implies that most of the incident photons will not be absorbed in the space charge region if the geometry of the photoanode is planar because the absorption depth (defined as $x^{-1} = \text{depth at which 63}\% \text{ of the photons are absorbed according to the Beer–Lambert law}$) for hematite ranges from 120 to 46 nm for photon wavelengths from 550 to 450 nm [17]. This would not be problematic if the diffusion length of the minority carriers was large.
compared to the absorption depth, but it is only 2–4 nm (challenge 5). The effect of these limitations emerges as low quantum efficiencies, especially for the longer wavelengths. This is shown in Fig. 4.3 (right) with the highly efficient Nb-doped single crystals [53]. Even though peak quantum efficiency is close to 30% at $\lambda = 370$ nm, it drops to below 5% at 450 nm and even though all of the incident irradiation is absorbed at 550 nm in the 0.1-mm thick films used, the quantum efficiency is less than 1%.

The above challenges and limitations suggest that the standard planar single crystal or sintered disk electrode geometries are not suitable for hematite, and indeed recent efforts with hematite have shown that challenges 3–5 can be overcome by optimizing the electrode morphology. These recent and exciting advances in hematite structuring are described in detail in the next section. In addition, the renewed efforts with structuring hematite have instigated further research to overcome challenges 1 and 2 listed above. These efforts are reviewed in Sect. 4.4.

### 4.3 Morphology Control of Fe$_2$O$_3$

Both the small (ca. 10 nm) space charge layer present in highly doped hematite (resulting from the necessity for conductivity) and the short (ca. 4 nm) diffusion length of minority carriers discussed in the previous section suggest that the ideal morphology of a hematite photoanode would be one where all of the hematite is within 10–20 nm of the semiconductor/liquid junction (SCLJ). This would give photogenerated holes a high probability of arriving at the SCLJ to oxidize water. Of course, one could simply deposit a 10 nm layer of hematite on a transparent substrate using a traditional deposition method. Then the poor absorptivity of such a film could be overcome by stacking multiple films in tandem. Indeed, this solution was first proposed by Itoh and Bockris in 1984 [55]. While this could fundamentally resolve the issue, it is cumbersome and expensive to implement, and in practice the thin iron oxide films were found to have poor performance due to the increased recombination of the photogenerated holes [54]. Fortunately, the recent development of tools to control the dimensions and morphology of materials at the nanometer length scale has offered more practical morphological solutions for hematite. Of course, the application of nanostructuring techniques to PEC solar hydrogen production has included more materials than simply hematite, with materials including WO$_3$ and TiO$_2$ also seeing noteworthy development. Two recent review articles from the research groups of both Augustynski [57] and Van de Krol [58] present a general overview of this topic. Currently, using new nanostructuring techniques, research groups are designing hematite with the proper morphology, feature size and dopants in order to enhance its ability to split water from solar irradiation. This section examines how directing the nanostructure of hematite has led to significant advances in its use as a material for photoelectrochemical water splitting. This is done by analyzing how different deposition techniques have allowed the morphology control of hematite and affected its performance as a photoanode.
4.3.1 Colloidal Hematite

The nanoscopic manipulation of hematite was first mastered by Matijević’s group in the early 1980s [59]. They created well-defined hematite colloidal particles through a hydothermal route and demonstrated morphological control by changing the concentration and the oxidative environment present during nucleation and growth. Spheres, disks, and plates of hematite were observed with lengths in the 100 nm range. However, no photochemistry was investigated with materials of this type until Moser and Grätzel reported halide photooxidation with colloidal dispersions of α-Fe₂O₃ 2 years later [60]. In this work, 60 nm particles dispersed in aqueous solution were irradiated with 347 nm laser light and the extremely rapid (~10 ns) formation of I₂⁻ from iodide at high quantum efficiency (>80%) was observed. This result suggested an advantage of employing nanosized hematite where a larger fraction of the photogenerated carriers are produced closer to the surface as compared to the single crystal hematite or sintered pellets which had been reported at the time. Similar studies following shortly thereafter used various electron relays (e.g., methyl viologen) and hole scavengers (e.g., methanol) to investigate the ability of Fe₂O₃ colloids to produce hydrogen from water [61–63]. While the conversion efficiency remained quite low even in these sacrificial systems, the optimum particle size, balancing bulk recombination and light harvesting, was estimated to be 100 nm in this work [63].

To date there have been numerous reports investigating the photocatalytic properties of colloidal α-Fe₂O₃. Most recently, layered “nanoflowers” [64] and hollow nanoparticles [65] synthesized via a simple solvothermal approach have shown enhanced photocatalytic activities for the degradation of Rhodamine B and diethyl phthalate, respectively, due to their large surface area. While these studies can give insight to the relative catalytic activity of different structures, given the band edge energies of hematite and the sluggish oxygen evolution reaction (OER) kinetics, it is likely that Fe₂O₃ colloids will not find application for water splitting but remain at the degradation of organics. However, the continued development of the precise control over hematite nanoparticle shape and size, like that recently demonstrated by Hu and Yu [66], should allow access to ideal electrode morphologies in the future.

4.3.2 Porous Thin Films from Colloids

Even if hematite nanoparticle dispersions could split water efficiently, the problems of gas separation and the reverse reactions (hydrogen oxidation and oxygen reduction) occurring in the heterogeneous mixture would remain. A more practical approach to water splitting is arguably a PEC cell with the anode and cathode evolving oxygen and hydrogen separately. This configuration provides the challenge of creating hematite with large surface area-to-volume ratio on the surface of
a conducting support. The most straightforward way to create nanostructured photoelectrodes of hematite is to coat a sol onto an electrode and sinter to enforce interparticle connections—a method first perfected for solar energy conversion with TiO$_2$ in the dye-sensitized solar cell. This method was first attempted with hematite by Bjorksten et al. in 1994 [67]. Here, Fe$_2$O$_3$ sols created by the hydrolysis of FeCl$_3$ were concentrated and combined with surfactant before they were doctor-bladed onto transparent conducting oxide substrates (F-doped SnO$_2$) and sintered at 560$^\circ$C. Micron-thick, porous thin films of necked hematite were observed to have good adhesion to the substrate and a primary particle size in the 25–75-nm range (see Fig. 4.4, left). However, the quantum efficiency (IPCE) of these photoanodes toward water splitting was quite low—on the order of 1% at 400 mV vs. SCE in 0.1 M NaOH (1.4 V vs. RHE) with 400 nm incident irradiation (see Fig. 4.4, right). The observations that the IPCE was 100 times lower when the anode was illuminated from the hematite/electrolyte interface as compared to the substrate/hematite interface and that the quantum efficiency did not improve significantly (only to 1.7%) when LiI was added as a hole scavenger led the authors to conclude that charge carrier recombination was the controlling factor for the photocurrent. The higher quantum efficiency of similar particles when dispersed in electrolyte [60] pointed to grain boundaries between the particles to be the cause of the recombination and poor performance. This limitation was addressed by Qain et al. [68] by altering the film thickness to optimize the light absorption/carrier transport issue. However, no significant improvement was obtained. Cathodic spikes observed during transient

![Fig. 4.4](image-url)  
Scanning electron micrograph (left) of the edge of a porous electrode made from sintered Fe$_2$O$_3$ sols on a conducting cassiterite (SnO$_2$) support (magnification 5,800 times). Absorption spectrum of the hematite electrode and quantum efficiency as a function of wavelength (right) obtained in 0.1 M NaOH at 1.4 V vs. RHE when illuminating the electrode through the substrate (SE) and directly onto the interface with the electrolyte (EE). Note that the scale of the IPCE values differs with a factor of 100 in the two cases. From [67] used with permission.
measurements in this work suggested that no space charge layer was present, and that electron recombination at the hematite/electrolyte interface was also an important limitation.

Recent work on Fe$_2$O$_3$ films prepared by colloidal methods has identified a possible application of sensing organics in gas streams due to a change in electrical resistivity [69, 70] and suggested that the inclusion of metal nanoparticles [71, 72] or the presence of a spinel phase (γ-Fe$_2$O$_3$) [73] may increase the catalytic activity of these electrodes. However, a fundamental issue of particle size recognized in the seminal work [67] has yet been unresolved: Despite attempts to use particle sizes commensurate with the hole diffusion length (<10 nm), the temperatures required to sinter and make the films photoelectrochemically active result in primary particle sized too large to expect high photocurrents. Incorporating dopant atoms [74] and controlling feature size with templating techniques [75, 76] may result in an increase in performance for hematite films prepared by this inexpensive and scalable technique.

### 4.3.3 Fe$_2$O$_3$ Nanowire Arrays

An elegant solution to the problem of grain boundaries in hematite films created from the colloidal approach is to use nanosized rod or wire arrays. An array of single crystalline nanorods with diameters in the 10-nm range, attached and oriented orthogonally to the conducting substrate would eliminate grain boundaries, enhancing conduction to the substrate while still allowing photoproduced holes to efficiently reach the semiconductor/electrolyte interface. A simple method to create hematite arrays on a variety of substrates via a metastable akaganeite phase produced from the controlled precipitation of Fe(III) in aqueous solution was first reported by Vayssieres et al. [77], and investigated for water photoelectrolysis soon after [19]. Bundles of individual 5 nm nanorods with an average diameter of 50 nm and a length of 0.1–1.5 μm were investigated in perpendicular and parallel orientations to the substrate (Fig. 4.5, left). While the authors were able to show a small improvement when controlling the orientation (5% IPCE at 360 nm for the perpendicular nanorods vs. 3% for the parallel) the overall photocurrents remained low under white light illumination, even with hole scavenging iodide present. The large difference between the electrode performance when illuminating from different sides of the semitransparent photoanode observed in this work [79], and a recent report [80] examining the surface photovoltage on electrodes prepared the same way, suggest that bulk defects and the slow oxygen evolution kinetics of hematite are the major factors limiting the performance of this system.

Despite the poor performance of the nanowire arrays prepared by this technique it remains an intriguing morphology for hematite photoanodes. Another facile method to produce hematite nanowires by the simple thermal oxidation of iron foils has been reported by many groups [78, 81–84]. Due to the increased volume of the oxide over the metal, when foils of iron are thermally oxidized under the right
conditions, arrays of Fe$_2$O$_3$ nanowires spontaneously grow from the surface of the iron. With diameters of 20–40 nm and lengths of up to 5 μm (see Fig. 4.5, right) these nanowire arrays would have large surface area, sufficient light absorption and a direct path for the conduction of electrons to the substrate (basal planes are oriented perpendicular to the substrate [83]), making them a very attractive morphology for hematite. This thermal oxidation method has been demonstrated on multiple (even transparent) substrates [84], and methods for doping the particles have been employed [78]. However, no realistic reports of activity toward water oxidation have appeared. This may be due to the defects present in nanowires prepared this way or the presence of suboxides near the interface of the substrate which greatly enhance recombination [85].

### 4.3.4 Electrochemical Fe$_2$O$_3$ Nanostructurin

The development of nanostructuring techniques using potentiostatic anodization has provided another possible route to create structured hematite photoelectrodes. While it has long been known that the potentiostatic anodization of certain metals can create structuring on small length scales, only recently have methods been developed to prepare ordered arrays of nanosized tubes or pillars with this technique. For example, ordered nanotube arrays of TiO$_2$ created using potentiostatic anodization have been successfully employed in multiple device applications including the solar photolysis of water [86]. Like the nanowire approach, these nanotube arrays have the potential to reduce crystal grain boundaries and enhance electron conduction to the current collector. In addition, the small thickness of the tube walls should allow for efficient hole collection. Recently, Grimes and coworkers showed that iron foils could be nanostructured using anodization in a glycerol-based electrolyte containing 1% NH$_4$F + 1% HF + 0.2% HNO$_3$ [87]. Ordered nanopores
were observed with pore size ranging from 20 to 250 nm and depths up to 600 nm depending on the anodization voltage and time (see Fig. 4.6, left). Under simulated solar illumination these photoanodes produced a photocurrent of 0.05 mA/cm² at 0.4 V vs. SCE in 1 M NaOH. This current increased tenfold with the addition of 0.5 M H₂O₂ indicating good electron transport in these films, and suggesting the surface kinetics of the water oxidation reaction were limiting the water-splitting photocurrent. The same research group improved on this result by alloying the iron with titanium before anodization [89]. Water-splitting photocurrents were obtained from films that were 6.6% Fe, based on metal, and were largely anatase TiO₂. They observed reasonable photoactivity with a quantum efficiency of ca. 4% with 400 nm light at 1.5 V vs. RHE. In contrast to this Fe–Ti–O work, very well-defined nanotube arrays (see Fig. 4.6, right) created from pure iron foils have been subsequently reported by a different research group [88]. In this work, a single anodization step with 0.1 M NH₄F + 3 vol% water in ethylene glycol created nanotubes with walls less than 50 nm and lengths of about 1.5 μm. After an optimized annealing treatment these electrodes were found to be a mixture of both hematite and maghemite by XRD and had only small photocurrents (160 μA/cm², 1.23 V vs. SCE).
RHE in AM 1.5 light compared to 120 μA/cm² dark current). However, a double step anodization procedure with the addition of sodium tripolyphosphate in the electrolyte, and the subsequent annealing created a dendrite-like morphology situated over the tube array. Electrodes prepared this way exhibited a photocurrent of about 1 mA/cm² at AM 1.5 and 1.23 V vs. RHE. This is quite remarkable given the presence of maghemite and as the films were not intentionally doped to increase the carrier concentration. However, a large carrier concentration \(10^{21} \text{ cm}^{-3}\) was found via Mott–Schottky indicating a large amount of dopant or impurity present. The authors attribute the high photocurrent to the double-layered structure which includes both a large surface area for water oxidation and a vertically orientated nanotube array for electron transport.

In subsequent work by the same group, a sono-electrochemical anodization method was employed to obtain either nanoporous or nanotubular Fe₂O₃ depending on the anodization conditions [90]. The optimization of conditions gave nanotubes with walls an impressive 5–6-nm thick and microns in length. Annealing these films resulted in pure hematite which gave similar photocurrents (ca. 1 mA/cm² at AM 1.5 and 1.23 V vs. RHE) and similar carrier densities to the previous work. While the initial tube size presented in this work is remarkable and an important step for obtaining hematite with the ideal morphology, the tubes were initially amorphous Fe₂O₃. The annealing step necessary for crystallinity and photoactivity probably demolishes the meticulous nanostructuring through a sintering process like has been observed for fine-structured hematite at high temperatures [67].

In addition to the nanostructuring of Fe₂O₃ with potentiostatic anodization, hematite photoanodes have been prepared by electrodeposition of precursors from solution. Recently, McFarland and coworkers reported an electrodeposition method of iron hydroxides from FeCl₃ solutions. The subsequent annealing at high temperature (700°C) then resulted in porous hematite films [20, 91]. This method readily allows for the incorporation of dopants which were found to have an effect on the morphology of the sintered electrodes. With feature sizes of 30–40 nm the observed IPCEs of 8% at 400 nm at 1.2 V vs. RHE in 1 M NaOH with 15% Mo doping are a notable result. However, the large particle size remains a limitation to higher photocurrents. Another example of electrodeposition has been recently reported by Spray and Choi [92] using an anodic method as opposed to the cathodic method investigated by the McFarland group. They were able to demonstrate impressive morphology control ranging from wire arrays to porous films by varying the solution pH. The films were photoactive in an electrolyte-containing iodide, but water oxidation photocurrents were not reported.

In general, it is clear that major advances have been made recently in the nanostructuring of hematite by electrochemical methods. However, the potential to oxidize water with hematite prepared in the methods described above has been limited by the crystallinity and ultimate feature size after the requisite annealing.
4.3.5 Spray Pyrolysis Techniques

The colloidal, nanowire, and anodization approaches to deliberately nanostructure hematite have provided great leaps in hematite nanostructuring but only moderate success so far for water-splitting photocurrents. However, from its first report [93] the spray pyrolysis (SP) method has continuously produced thin films of photoactive hematite. Typically, an aqueous or ethanolic solution of Fe\(^{3+}\) is simply sprayed onto a hot (~400°C) substrate and photocurrents in the 0.5–1.0 mA/cm\(^2\) range can be obtained under simulated solar illumination (1 M NaOH, 1.6 V vs. RHE) [94–97] with the exception of one report [98] that is improbable in view of the action spectrum reported by the authors [99]. Surprisingly, modest photocurrents are obtained with this method even without intentionally adding dopants to increase the conductivity of the Fe\(_2\)O\(_3\). However, it is likely that the commonly used precursors like FeCl\(_3\) or Fe(AcAc)\(_3\) leave chlorine or carbon residues behind that could act as electron donors. Typically, this technique can produce films with nanometer crystalline domain size, but the films are generally compact (not porous) and the surface area limits the conversion efficiency. However, the ease and reproducibility of the technique allow for fundamental studies on dopants [95, 100] and substrate effects [96], and careful optimization has produced films with a notable quantum efficiency (IPCE = 25% at 400 nm and 1.42 V vs. RHE) [100].

In a recent report from Duret and Grätzel, the spray pyrolysis technique was modified to create nanostructured films [94]. Here an ultrasonic spray nozzle was used to create very fine droplets of the precursor which were entrained in a flow of dry air and carried rapidly and horizontally over a heated substrate. The low concentration of the precursor in the gas stream allowed the slow growth (~100 nm/h) of mesoscopic hematite with a leaflet-type structure consisting of 100-nm-sized platelets with thickness of 5–10 nm bundled into 50 nm sheets oriented perpendicular to the F-doped SnO\(_2\) support (see Fig. 4.7, top). These ultrasonic spray pyrolysis (USP) films were found to perform impressively for water splitting, producing 1.3 mA/cm\(^2\) under AM 1.5 (100 mW/cm\(^2\)) simulated solar irradiation in 1 M NaOH with an applied potential of 1.23 V vs. RHE. The IPCE was found to be 16% with 375 nm illumination at the same applied potential, and over 30% at 1.6 V vs. RHE. This was a threefold increase over control films created by a conventional spray pyrolysis method in the same work, and even an improvement over optimized photoanodes created by SP [100]. The advantage of the nanostructured morphology could be directly observed by comparing normalized IPCEs of the UPS and SP electrodes (Fig. 4.7, bottom). The IPCE is higher for longer wavelengths (400–600 nm), indicating the more efficient conversion of holes due to the morphology. Subsequent work on these films revealed the importance of Si doping, and showed an enhanced (110) reflection from XRD as compared to powder samples indicating a preferred orientation of the (001) basal plane normal to the substrate [101, 102]. The vertical orientation of the (001) planes is thought to be very important to the conduction of photogenerated electrons to the conducting substrates, as this is the least resistive direction of conduction in hematite.
However, because the larger surface of the platelets is the basal plane, this morphology also requires the holes to be conducted across oxygen layers. This detail and the poor kinetics of water oxidation were thought to be limiting the performance in this system.
Continuing efforts from Grätzel and coworkers to perfect the nanometer length scale morphology of hematite thin films for water splitting beyond the USP films have led to the development of a simple, yet effective process to make nanostructured films of hematite found to be highly active for water photooxidation (Fig. 4.8). This method is based on the thermal decomposition of iron pentacarbonyl in an atmospheric pressure chemical vapor deposition (APCVD) reactor. Fe(CO)$_5$ vapors are transported, along with tetraethyl orthosilicate (TEOS) as a source of silicon dopant, by an air stream vertically to a substrate maintained at 420°C. This high temperature and the poor thermal stability of the Fe(CO)$_5$ (the half-life for decomposition at 300°C is 5.3 ms [103]) causes the homogenous nucleation of nanoparticles in the gas phase [104]. These particles are then subjected to a thermophoretic force which limits their approach to the hot substrate and results in fractal-like cauliflower structures suggesting a diffusion-limited aggregation (rather than surface reaction limited) mechanism [13]. Similar to the films prepared by the USP method, the silicon dopant was found to affect the

![Fig. 4.8 Morphology of Si-doped hematite films created by the APCVD technique as observed by SEM. The left side shows a cross-sectional view of the Fe$_2$O$_3$, prepared at a substrate temperature of 420°C (heater setpoint of 470°C), on top of the F:SnO$_2$ substrate. The right shows top-down views of films prepared with heater set points of 450°C (above) and 490°C (below). From [13] used with permission](image)
nanostructure, producing a smaller feature size in both cases and an alignment of the basal crystal plane normal to the substrate [102, 105]. While this latter observation is considered to be a major factor improving the performance of these electrodes by facilitating electron transport to the collector, another report found the crystalline orientation of hematite unimportant to the performance [106].

The optimization of the APCVD deposition technique led to films producing a photocurrent of 1.8 mA/cm² at 1.23 V vs. RHE under simulated solar illumination (AM 1.5, 100 mW/cm²) in 1 M NaOH. This was further increased to 2.2 mA/cm² by treating the surface with Co²⁺ ions, which act to lower the energy barrier for water oxidation and reduce the applied potential needed [105] (see also Sect. 4.4.2). With IPCEs of 42% at λ = 370 nm and 1.23 V vs. RHE these electrodes represent the most efficient Fe₂O₃ photoanodes for water oxidation to date – surpassing even the single-wavelength quantum efficiencies of Nb-doped α-Fe₂O₃ single crystals (IPCE = 37% at 370 nm and 1.23 V vs. RHE) [53]. This benchmark performance corresponds to 3.3% STH conversion efficiency when used in a tandem configuration (see Sect. 4.4.1). The PEC characterization of this champion electrode is shown in Fig. 4.9. Particularly notable is the increased IPCE of the photoanodes at wavelengths between 500 and 600 nm as compared to the Fe₂O₃ film prepared by the USP method [94] and especially the single crystal electrodes shown in Fig. 4.3. Because the high surface area and small feature size of the APCVD films the photocurrent action spectrum is much broader than the single crystal results: photons with longer wavelength have a higher chance for being absorbed within the hole diffusion length, resulting in much higher solar photocurrents than the single crystal electrodes.

Fig. 4.9 Photoelectrochemical characterization of Si-doped Fe₂O₃ electrodes prepared by the APCVD technique. The left graph shows the current-voltage characteristics in darkness and under simulated sunlight (100 mW/cm²) at a scan rate of 50 mV/s. (a) Unmodified Fe₂O₃; (b) the same electrode after cobalt treatment. The right graph shows the incident photon-to-current efficiency (IPCE) spectrum of the same electrode at different conditions. (a) Unmodified Fe₂O₃ at 1.23 V_RHE; (b) after cobalt deposition, at 1.23 V_RHE; (c) at 1.43 V_RHE; (d) at 1.03 V_RHE; (e) at 1.23 V_RHE but illuminated through the SnO₂ substrate; (f) absorption spectrum of this electrode. From [105] used with permission.
In continuing work with this deposition technique Cesar et al. [13] examined the deposition mechanism, and showed a large temperature effect on the nanostructure which corresponded to the performance of the electrodes. In addition, the electrodes were studied by impedance spectroscopy; a Mott–Schottky analysis gave the calculated donor density of the Si-doped Fe$_2$O$_3$ films as $1.7 \times 10^{20}$ cm$^{-3}$ – close to the predicted $6 \times 10^{20}$ cm$^{-3}$ if all of the 1.6 at% Si [105] was substitutionally incorporated into the lattice as Si$^{4+}$. This observation suggests that a space charge layer of ~5 nm could be present and aiding hole transport the SCLJ even in the small features on the top of the APCVD film. The high donor density also created good electron conductivity and films up to 1 µm delivered similar photocurrents as thinner films when illuminated from the semiconductor/electrolyte interface even though photons were absorbed further from the conducting substrates. This suggests that electron percolation through the film is not a limiting factor.

Recently another group has reported the application of the APCVD technique for Fe$_2$O$_3$ photoanodes using both ferrocene and iron pentacarbonyl as iron sources, and different sources of Si for doping [107, 108]. First, ferrocene as a precursor was shown to be effective in preparing photoactive Fe$_2$O$_3$ films without a separate dopant species [108]. Photocurrents were ca. 0.6 mA/cm$^2$ at 1.23 V vs. RHE in 1 M NaOH under simulated solar irradiation. Furthermore, the larger feature size observed in the ferrocene-prepared electrodes (50–100 nm) as compared to the Fe (CO)$_5$ indicates the importance of the decomposition of the carbonyl precursor and the nucleation of nanoparticles before reaching the hot substrate. Next, the silicon doping was studied by varying the Si precursor species and feed rate [107]. In addition to TEOS, tetramethyl orthosilicate (TMOS) was also found to be an effective source of Si. XPS studies verified the +4 oxidation state of Si in both cases. In addition, trimethyl orthoformate (TMOF) was used as a silicon-free control, and resulted in electrodes with significantly less photocurrent. Although their findings did not result in a definitive explanation of the silicon’s role, it did suggest that a combination of factors including the formation of lattice defects and nanoparticle cluster formation caused by the Si precursor’s decomposition and incorporation, additional to the classical doping effect, may be contributing to the high performance observed in these electrodes.

Despite the impressive performance of the APCVD iron oxide films the solar photo current remains low compared to the theoretical value of 12.6 mA/cm$^2$ predicted for a material with this bandgap [6]. This is because, despite absorbing almost all of the light (see Fig. 4.9, left, curve f) the quantum conversion efficiencies are relatively low and are especially poor (<20%) in the region where hematite has an indirect bandgap transition ($\lambda = 610–450$ nm). Importantly, there is over 10 mA/cm$^2$ of solar photocurrent available in this wavelength range. The poor performance can be primarily attributed to the cauliflower-type morphology. While this morphology has an appropriate feature size of 10 nm at the SCLJ, the individual cauliflower structures consist of thick stems which increase the number of photons absorbed far from the SCLJ.
4.3.7 Extremely Thin Absorber–Host-Scaffold Approach

In an extension of the stacked thin film approach proposed by Itoh and Bockris [55] an increased photoresponse should also be obtainable with hematite by coating a thin film conformally on a suitable nanostructured collector with a large specific surface area, in analogy to the dye-sensitized solar cell (DSC) [109] and similar extremely thin absorber (ETA) photovoltaics [110]. In the case of the DSC, the photoanode is constructed from a light-absorbing dye molecule anchored to a high surface area mesoporous TiO$_2$ scaffold, which collects and transports the photoexcited electrons. The DSC’s host-scaffold/guest-absorber approach effectively decouples light harvesting and charge transport while maximizing the IPCE by means of the thin (mono)layer of absorber on the high surface area, transparent collector. Given the success of this concept for photovoltaic energy conversion as well as the light absorbing and photogenerated hole transport limitations of $\alpha$-Fe$_2$O$_3$, this is a promising direction for increasing the quantum efficiency of hematite.

A difficulty in applying this method for hematite is finding a suitable host material which satisfies the following criteria: First, the conduction band of the scaffold material must lie lower in energy than the conduction band of hematite to allow efficient electron transport across the host/guest interface and second, the scaffold material should have a larger band gap than hematite so as to not compete with the light absorption. Few oxide semiconductors meet both of these requirements. However, Grätzel and coworkers have recently demonstrated this concept for hematite photoanodes using nanostructured WO$_3$ as the host scaffold [8]. In this work the host scaffold was coated with a thin layer (ca. 60 nm) of Fe$_2$O$_3$ by the APCVD method (see Fig. 4.10, top) and a 20% increase in the photocurrent was observed over control samples prepared on the flat F:SnO$_2$ substrates. This increase was attributed to an increase in the absorbed photon-to-current efficiency (APCE), especially in the wavelength range from 500–600 nm (Fig. 4.10, bottom).

While this demonstration of the host/guest approach indicates it is a promising and accessible methodology to fully exploit the light harvesting capabilities of hematite for solar water splitting, the best results were obtained with iron oxide coatings 60 nm thick over the host scaffold. This is too thick based on the hole transport limitations already discussed. Attempts to use thinner films of iron oxide resulted in poor performance, similar to other reports of thin films Fe$_2$O$_3$ [54]. Further understanding and control of the interface between the iron oxide and the conducting substrate as well as the electronic properties and crystallinity of ultra-thin films of Fe$_2$O$_3$ are necessary to advance this promising nanostructuring approach.
Fig. 4.10  (Top) Scanning electron micrographs of the WO$_3$/Fe$_2$O$_3$ host/guest system. Image (a) shows the morphology of the WO$_3$ host scaffold before iron oxide deposition. After 1 min of Fe$_2$O$_3$ deposition at the standard deposition rate (~100 nm/min) the WO$_3$ film is covered in iron oxide nanoparticles (c) while at a lower deposition rate (~10 nm/min, edge of growth spot) the iron oxide nanoparticles are seen just starting to cover the WO$_3$ scaffold (b). The cross-section of the film, corresponding to (c), is shown in (d). (Bottom) The APCE (solid lines) at 1.43 V$_{RHE}$ and the transmittance, $T$ (broken lines) of photoanodes prepared with a 1-min Fe$_2$O$_3$ deposition time plotted against the incident light wavelength, $\lambda$. The host/guest and the control electrode responses are indicated by dark and light lines, respectively.
Many different approaches have been taken to nanostructure Fe$_2$O$_3$ for use in PEC water splitting, and at least as many methods of characterizing the electrodes have also appeared. As such it can be difficult to compare the results in the literature, especially when variations in light sources give large over-estimations of photocurrents due to superfluous UV light. A recent and important contribution by Murphy et al. addresses these issues and suggests reliable methods for evaluating electrode conversion efficiencies [6]. While not all of the reports discussed in this section have followed these suggestions it can still be helpful to compare the top performances from the different structuring techniques and the resulting performance. This is done in Table 4.1.

It is clear, by looking at Table 4.1, that while many methods of structuring hematite have been reported, the IPCE values at ca. 400 nm rarely come close that of the single crystal or sintered polycrystalline electrodes prepared during the first
decade of work. Even the champion electrodes made by APCVD merely match the IPCE of the single crystal work at ca. 400 nm without the enhanced surface treatment. The key to the increased photocurrents in the APCVD electrodes has, of course, been the nanostructuring, which allows for the more efficient conversion of the longer wavelengths of light. Further work on the nanostructuring will, no doubt, see solar photocurrents rise, but the increase in IPCE beyond 40% at 400 nm and 1.23 VRHE will require research to tackle hematite’s other remaining challenges.

4.4 Overcoming Hematite’s Remaining Challenges

Many techniques have been used to nanostructure hematite, and while clear advances have been made over planar, single crystal approaches, in general the performance of nanostructured electrodes has remained poor compared to the quantum efficiencies found in optimized photoanodes prepared from TiO2 or WO3. While this is partly due to the absence of a method to prepare a “perfect” nanostructure, there are still limitations in (1) the understanding of the critical material properties affecting efficiencies and (2) the surface oxidation reaction kinetics. In addition, based on the flatband potential of hematite it cannot be used to split water unassisted by a bias voltage. In this section, these limitations are clearly detailed and new promising approaches to overcome these remaining challenges are reviewed.

4.4.1 Flatband Potential and Tandem Cell Approach

As mentioned in Sect. 4.2.5, one of the major drawbacks of hematite is that its flat band potential is below the water reduction potential. This means that the electrons exiting the conduction band of hematite and traveling to the cathode do not have sufficient energy to reduce water and create hydrogen; a small external bias must be added to hematite to complete the water-splitting reaction. The overall efficiency of solar water splitting then must account for the power input by this external bias. However, a clever solution to this problem was invented by Augusttynski and Grätzel [111]. Since hematite only absorbs solar irradiation up to wavelengths of 600 nm, transmitted photons (λ > 600 nm) can be converted by another photosystem to provide the necessary potential to complete the water-splitting reaction and accomplish unassisted solar hydrogen production from water. This second photosystem is most practically a photovoltaic device as photocathode materials typically suffer from stability issues. A promising candidate for this photovoltaic device is the DSC as these devices can be produced for one-third to one-fifth the cost of a conventional silicon solar cell and the absorption range can be tuned by selecting the appropriate dye.
Ideally in this hematite/DSC water-splitting tandem cell, only one photovoltaic system would provide the necessary potential to complete the water-splitting reaction. However, due to the large overpotential needed for water oxidation (see Sect. 4.4.2) and other loss processes, two solar cells are needed to provide enough voltage to the current state-of-the-art Fe₂O₃. For example, saturation (plateau) photocurrents from Fe₂O₃ prepared from APCVD are not reached until 1.4 V vs. RHE. While a traditional solution can still be envisioned with two solar cells with half the active area each behind the hematite photoanode, the recent development of new dyes for the DSC opens up new possibilities for optimizing light harvesting in these water-splitting tandem cells. Different architectures for the hematite/DSC system were recently investigated by Brillet et al. [112] and it was found that a three-level tandem (hematite/suqarane dye DSC/black dye DSC) was the most efficient at splitting water with a calculated overall efficiency of 1.4% solar to hydrogen. This is much lower than the predicted 3.3% with this system based on iron oxide photocurrent alone, and losses were attributed to light scattering and the large overpotential for water oxidation by hematite.

A tandem cell configuration would not be needed if the flat band potential of hematite could be shifted so that water could be directly split by solar light. While numerous groups have reported a $V_{fb}$ 0.3 – 0.4 V lower than the hydrogen evolution potential [14], Hu et al. recently suggested that a CoF₃ treatment of Ti-doped Fe₂O₃ could actually shift the flat band potential and allow the direct hydrogen production by hematite [113]. While additional investigation of this claim is necessary to verify sustained hydrogen production at zero bias, it is indeed a promising route to reduce or eliminate the necessary applied bias.

4.4.2 Oxygen Evolution Overpotential and Surface Catalysis

Based on the flat band potential of hematite an external bias of only 0.3–0.4 V vs. RHE should be necessary to initiate the water-splitting reaction. Once the applied bias is greater than the flatband potential, the band bending will drive photogenerated holes to the surface for reaction. However, photocurrent onset potentials are usually not observed until 0.8–1.0 V vs. RHE even for the single crystal electrodes. The remaining potential of ca. 0.5 V is due to the overpotential for water splitting caused by unfavorable surface properties. These unfavorable properties have two origins. First, there is evidence that oxygen vacancies [17, 23, 46] and crystalline disorder [13] can be present at the surface. This would create additional Fe²⁺ recombination centers for holes, and can even result in Fermi level pinning in some electrodes [94]. Second, the reaction kinetics for water oxidation are slow, not just due to the complicated four-electron mechanism that must occur, but also compared to other oxide semiconductors [79]. This may be due to the more pronounced metal 3d character of the valence band compared to other oxides [23].
The first surface challenge can be addressed by strong oxidation conditions [46] and careful hematite preparation, but the slow water oxidation kinetics are probably intrinsic to hematite. Nevertheless, methods have recently been found to increase the oxidation rate and thus reduce the overpotential. For example, the water oxidation by cobalt has been extensively studied and is known to be particularly rapid [114]. Indeed the treatment of Fe₂O₃ photoanodes (prepared by APCVD) with a monolayer of Co²⁺ resulted in a ca. 0.1 V reduction of the photocurrent onset potential [105]. Since this treatment also increased the plateau photocurrent it is good evidence that the reaction rate was increased, and the Co²⁺ did not just fill surface traps. Following the report of a remarkably effective cobalt-phosphate (Co-Pi)-based water oxidation catalyst [115], the overpotential was reduced even further on hematite photoanodes by Gamelin and coworkers [116]. Their results are shown in Fig. 4.11.

The application of the Co-Pi catalyst resulting in a shift of the onset potential by 0.25 V. The effect is further supported by the increase in IPCE seen in Fig. 4.11. Applying this treatment to the high performance Fe₂O₃ anodes could possibly result in photocurrents nearing 3 mA/cm² at 1.23 V vs. RHE. Further work is necessary to understand the effect of this catalyst over the standard Co²⁺ treatment, and to better characterize the reaction mechanism occurring at the surface of both treated and untreated hematite. Interestingly, recent reports describe a very complex anisotropic reactivity of hematite at the SCLJ due to the anisotropic crystal structure and exposed crystalline planes [117, 118]. Fully exploiting the most reactive crystal face of hematite might further reduce the required overpotential for water oxidation.
4.4.3 Continued Material Considerations

To conclude, it is clear there is a strong dependence of the performance of synthetic hematite on the deposition technique. While methods such as spray pyrolysis and CVD consistently produce electrodes photoactive for water oxidation, solution-based methods such as sol–gel approaches have failed to produce especially photoactive hematite. This is certainly related to the quality of the prepared material in terms of crystallinity and impurity concentrations. Aqueous methods of preparing hematite typically pass through a phase containing iron hydroxide (e.g., akaganeite, lepidocrocite, or goethite) but primarily hematite is detected after at annealing at 500°C. However, it has been shown that at temperatures up to 800°C, a nonstoichiometric composition remains in hematite when prepared in this way [119]. This results in a distortion in the crystal structure due to cation vacancies and remaining hydroxides. A structural temperature dependence has also been shown to affect the color of hematite due to the subtle relaxation of the iron positions in the crystal lattice and the resulting change in the optical reflectivity [120]. This was observed to occur at similar temperatures of 800°C. In addition, nanoparticles of hematite prepared via an aqueous precipitation have shown a size and temperature-dependent concentration of spinel defects at the surface. The formation of these defects on hematite was shown to be governed by the thermodynamics of the α–γ phase transition with decreasing particle size [121]. Raman spectroscopy has shown that similar defects on the surface of highly efficient anodes prepared by APCVD are important to the performance [13].

The above factors can all play a major role in both surface and bulk recombination, and thus care must be taken to prepare and characterize hematite for PEC application. Empirically, these problems can be avoided by simply heating the electrode to sufficiently high temperatures as was done with the sintered polycrystalline work [28] or the recent work by McFarland and coworkers [20, 91]. However, there is a clear tradeoff between phase purity and feature size as at high temperatures rapid sintering of nanostructures occurs. In general, the important material parameters needed to produce highly photoactive hematite must be clarified in order to be able to apply the most appropriate nanostructuring technique and obtain an ideal morphology while still ensuring photoactivity.

In addition to the structural properties, there is an unmistakable need to increase understanding of the hematite–substrate interface. A photoinactive “dead layer” has been observed in hematite near the surface of the supporting electrode [8, 13, 54, 105] which can be ameliorated with the addition of an ultra-thin Si-based insulating layer [105]. While fundamental studies of this “dead layer” have increased understanding and further reduced its effect [96], the ability to create ultra-thin films (~20 nm) of hematite with quantum efficiencies (APCEs) matching the single crystal or Si-doped APCVD electrodes has not yet been demonstrated. A fundamental understanding of hematite/conductor interfaces is needed to overcome this issue which currently limits its application to ETA configurations.
Finally, the low quantum efficiencies (even at short wavelengths in optimized single crystal photoanodes), the ultrafast carrier recombination, and the interbandgap states found in hematite suggest that this material possesses intrinsic limitations that will be difficult, if not impossible, to resolve given our lack of understanding. However, recent combinatorial efforts to optimize both the alloy composition [122, 123] and doping in hematite [124] suggest that some improvement can be made of the standard “single dopant” approach. For example, the proper doping with multiple atoms may increase the hole and electron mobilities by changing the dielectric constant of the crystalline phase. Continued efforts along these lines and further fundamental studies of the PEC properties of hematite are necessary, parallel with the nanostructuring efforts, to completely harness the potential of hematite for solar water splitting. Indeed, since the initial writing of this chapter many important advances already have been reported in the performance and understanding of hematite [125] and many more are sure to follow.

References


5.1 Introduction

An efficient solar energy conversion system for hydrogen production from water, referred to as a solar–hydrogen system, is most important for genuinely sustainable development. However, for such a system to be practical and economical, we must address several remaining issues related to solar-to-H$_2$ efficiency, cost, and stability.

Solar–hydrogen systems using semiconductors fall mostly into the following three categories: solar-cell–electrolysis systems, semiconductor–photoelectrode systems, and photocatalyst systems. As Fig. 5.1 shows, each has advantages and disadvantages. The solar-cell–electrolysis system is efficient but expensive; the photocatalyst system is simple and relatively inexpensive, but inefficient; and the semiconductor–photoelectrode system falls somewhere between the two in terms of efficiency and cost. The semiconductor–photoelectrode and photocatalyst systems have been developed in different fields and in different groups. The photoelectrode system has been developed at various places in the world, and the photocatalyst powder system has been developed mainly in Japan [1]. However, as Fig. 5.2 shows, their water-splitting reaction mechanisms are very similar. Therefore, interaction and exchange of information between researchers in both fields are highly desirable. In particular, screening of oxide semiconductor materials with high efficiency under visible light irradiation is a main issue for both fields.

In this chapter, we first discuss mixed metal oxide semiconductors used in both fields. We then discuss photoelectrochemical properties of BiVO$_4$ semiconductors. Finally, we discuss our high-throughput screening system for mixed metal oxide semiconductors.

---

K. Sayama
Energy Technology Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Central 5, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8565, Japan
e-mail: k.sayama@aist.go.jp

R. van de Krol and M. Grätzel (eds.), Photoelectrochemical Hydrogen Production, Electronic Materials: Science & Technology 102, DOI 10.1007/978-1-4614-1380-6_5,
© Springer Science+Business Media, LLC 2012
5.2 Mixed Metal Oxide Semiconductors Used for Photoelectrodes and Photocatalysts

Since the report of the Honda–Fujishima effect using TiO₂ under UV light [2], photoelectrodes using metal oxide semiconductors have been widely investigated for application to solar hydrogen systems. For increased efficiency of solar energy conversion, effective utilization of visible light over a wide wavelength region is necessary. It is noted that metal oxide semiconductors offer many advantages: they are generally stable, suitable for O₂ evolution, and produced easily compared to other types of semiconductor. Moreover, most oxide semiconductors are n-type and H₂ gas collection at the counter electrode is easier than that on the wide-area photoelectrode of p-type semiconductors. However, efficiencies of various conventional-type metal oxide semiconductor electrodes, such as single crystals or sintered pellets, are very low in the visible light region, even with the application of high external bias.
An excellent incident-photon-to-current conversion efficiency (IPCE >75%) for water decomposition under visible light irradiation was reported using a nanocrystalline WO₃ thin film on a conducting glass electrode [3, 4]; the film was prepared by spreading a colloidal solution of WO₃ on the conducting glass and firing it. Relatively high IPCEs were also reported for porous Fe₂O₃ thin films on conducting glass electrodes prepared by spraying the glass with a solution of iron salts [5–8]. It is worth mentioning that these high IPCEs were obtained with simple wet-chemical processes, although many defects in the porous film electrodes may be present. Fig. 5.3 shows the porous structure of these nanocrystalline electrodes. Porous thin-film photoelectrodes are advantageous in that the diffusion length of the holes in the valence band formed by band-gap photoexcitation is significantly shorter compared to the length in conventional thin film-based photoelectrodes. This is because the electrolyte solution can penetrate the pores of the entire nanocrystalline semiconductor film, and holes in the nanoparticles can reach the semiconductor–electrolyte interface in a short time. Porous thin-film photoelectrodes have a thickness of a few micrometers or less, whereas conventional photoelectrodes are a few hundred micrometers thick or more. Furthermore, the diffusion path length of the electron is shorter when light is irradiated through the conducting glass. Therefore, this special structure probably accounts for the high IPCE.

Charge separation occurs in the nanoparticles for porous photoelectrodes and in the space charge layer for conventional thick photoelectrodes. The mechanism of charge separation for porous photoelectrodes is similar to that for photocatalysts, because charge separation for powders of the latter also occurs in small particles. Therefore, porous photoelectrodes are sometimes called photocatalyst electrodes.

Table 5.1 lists some essential characteristics of semiconductors for photocatalysts and porous photoelectrodes. Common essential characteristics of the two include stability during illumination, a valence band potential more positive than 1.23 V (vs. RHE), and O₂ evolution capability on the surface. For photocatalytic water splitting, the bottom of the semiconductor’s conduction band potential must be <0 V (vs. RHE) for H₂ evolution, and loading with a co-catalyst such as Pt or NiOₓ is often needed for H₂ evolution. In contrast, for evolution of O₂ gas at a photoelectrode, a high mobility of electrons to the back contact via the semiconductor particles is essential, so
an n-type semiconductor is preferable. Even if the conduction band potential of the photoelectrode is more positive than 0 V vs. RHE, hydrogen evolution at the counter electrode can be readily achieved by applying an external bias potential. Therefore, one can consider the requirements on the band edge positions of photocatalysts to be more restrictive than those of photoelectrodes.

Semiconductor evaluation and screening procedures vary for photocatalysts and photoelectrodes. For photoelectrodes, a semiconductor thin film is prepared on a conducting substrate, and its photocurrent–potential dependence is measured easily by potentiostat, even if the photocurrent is very small. However, quantitative measurement of O2 gas evolution on photoelectrodes is not generally conducted, therefore, it is important to consider the possibility of oxidation reactions other than O2 evolution. For photocatalysts, semiconductor powder is suspended in aqueous solution, and evolved H2 and O2 gases under light irradiation are measured by on-line gas chromatograph. When the O2 evolution capability of the semiconductor powder is evaluated, a strong electron scavenger (oxidative sacrificial reagent) such as Ag+ ion is added to the reaction solution. Preparation of thin films, as is required for photoelectrodes, is not necessary. There are wide variations in the available methods to prepare a single phase of semiconductor powder. Therefore, photocatalysis is an effective method for screening single-phase semiconductor powders of mixed metal oxides. However, the disadvantages include requirements for large glass-made apparatus without air leakage and lengthy times needed for O2 gas analysis.

There have been some investigations on porous photoelectrodes using visible light active semiconductors which were developed in the photocatalysis field, shown in Table 5.2. WO3 is a typical simple oxide photocatalyst that evolves O2 using electron scavengers under visible light irradiation, and the porous WO3 photoelectrode has the highest IPCE in the visible light region. Bi-based mixed-oxide semiconductors are often used as both photocatalysts and porous photoelectrodes because of their special band structures. For conventional simple oxide semiconductors, the valence band consists mainly of O-2p orbitals if the metal d orbital is unoccupied, and the potential of the O-2p orbital is about +3 V vs. RHE (Fig. 5.4a). Since the required external bias for the photoelectrode increases

### Table 5.1 Required characteristics of semiconductors for photocatalysts and porous photoelectrodes

<table>
<thead>
<tr>
<th>Common points</th>
<th>Photocatalysts (for water splitting)</th>
<th>Porous photoelectrode (for O2 evolution)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stability during photoreaction</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Positive potential of valence band top (&gt;1.23 V vs. RHE)</td>
<td>n-Type semiconductor with high mobility of e&lt;sup&gt;−&lt;/sup&gt; to back contact</td>
</tr>
<tr>
<td></td>
<td>O2-evolution capability on the surface</td>
<td></td>
</tr>
<tr>
<td>Different points</td>
<td>Negative potential of conduction band bottom (&lt;0 V vs. RHE)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H2-evolution capability on the surface (with co-catalyst)</td>
<td></td>
</tr>
</tbody>
</table>
Table 5.2  Semiconductors investigated for water decomposition into H₂ or O₂ by photocatalysts and porous photoelectrodes under visible light irradiation

<table>
<thead>
<tr>
<th>Semiconductor</th>
<th>Bandgap</th>
<th>Photoelectrolysis</th>
<th>Reference</th>
<th>Photocatalysis</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>WO₃</td>
<td>2.8</td>
<td>IPCE = 75% (400 nm), $V_{op} = +0.5$ (RHE) 2.2 mA/cm² (1.23 V, RHE, 1 sun)</td>
<td>[3]</td>
<td>$O_2(Ag^+)$: 65 μmol/h (&gt;420 nm) QE = 2% (405 nm)</td>
<td>[1, 16–18]</td>
</tr>
<tr>
<td>BiVO₄</td>
<td>2.4</td>
<td>IPCE = 44% (420 nm, +1 V, Ag/AgCl, Na₂SO₄) $V_{op} = -0.2$–0.05 V (ibid)</td>
<td>[9]</td>
<td>$O_2(Ag^+)$: 421 μmol/h (300 W–Xe, &gt;420 nm) QE = 9% (450 nm)</td>
<td>[19]</td>
</tr>
<tr>
<td>Bi₂WO₆</td>
<td>2.8</td>
<td>IPCE = 0.1% (400 nm), $V_{op} = 0$ V (SCE, Na₂SO₄)</td>
<td>[10]</td>
<td>$O_2(Ag^+)$: 3 μmol/h (300 W–Xe, &gt;420 nm)</td>
<td>[20]</td>
</tr>
<tr>
<td>Bi₂MoO₆</td>
<td>2.7</td>
<td>$V_{op} = 0.15$ V (RHE), 3 μA/cm² (300 W–Xe)</td>
<td>[11]</td>
<td>$O_2(Ag^+)$: 55 μmol/h (300 W–Xe, &gt;420 nm)</td>
<td>[21]</td>
</tr>
<tr>
<td>BiZn₂VO₆</td>
<td>2.4</td>
<td>IPCE = 0.4% (420 nm, +1 V, Ag/AgCl, Na₂SO₄)</td>
<td>[12]</td>
<td>$O_2(Ag^+)$: 15 μmol/h (350 W–Xe, &gt;420 nm)</td>
<td>[22, 23]</td>
</tr>
<tr>
<td>BiFeO₃</td>
<td>2.1</td>
<td>$V_{op} = 0.6$ V (SCE), 5 μA/cm² (1 V, SCE, 500 W–Hg)</td>
<td>[13]</td>
<td>$O_2(FeCl_3)$: 1 μmol/h (500 W–Hg, &gt;420 nm)</td>
<td>[13]</td>
</tr>
<tr>
<td>Fe₂TiO₅</td>
<td>2.3</td>
<td>IPCE = 4% (400 nm), $V_{op} = +0.3$ V (SCE, pH = 7) 2 mA/cm² (at +1 V, SCE, 500 W–Xe)</td>
<td>[14]</td>
<td>Acetaldehyde degradation (UV): FeTiO₃→Fe₂TiO₅</td>
<td>[24]</td>
</tr>
<tr>
<td>TaON</td>
<td>2.5</td>
<td>20 μA/cm² (1 V bias, 300 W–Xe, &gt;420 nm)</td>
<td>[15]</td>
<td>$O_2(Ag^+)$: 380 μmol/h (&gt;420 nm), QE = 34%</td>
<td>[25, 26]</td>
</tr>
</tbody>
</table>
with a positive shift of the conduction band, the preferred method to decrease the bandgap is via a negative shift of the valence band potential. For Bi mixed-oxide semiconductors, the valence band potential is shifted negatively when hybrid orbitals of O-2p and Bi-6s are formed due to repulsion of the two orbitals (Fig. 5.4b). Bi mixed oxides are relatively stable, and so Bi-based mixed-oxide semiconductors have been investigated for use as photoelectrodes as well as photocatalysts. Among Bi-based mixed-oxide semiconductors, BiVO₄, shows the highest photocatalytic activity for O₂ evolution and the highest IPCE under visible light irradiation [9]. The IPCEs of other Bi-based mixed-oxide semiconductors are very low, but will probably improve if preparation of porous semiconductor film is optimized.

The valence band potential of TaON can also be shifted negatively, due to hybridization of the O-2p and N-2p orbitals. As a result, TaON powders show good photocatalytic activity for not only O₂ evolution but also H₂ evolution [25, 26]. The potential of the conduction band bottom is probably close to the potential of H⁺/H₂. Therefore, TaON is a promising semiconductor for photoelectrodes with the application of small external bias, although its photoelectrochemical performance is still poor [15].

There are many reports on Fe-based semiconductor photoelectrodes [5–8], but very few on O₂ evolution over Fe-based mixed-oxide photocatalysts. It is speculated that the reduction capability of electron scavengers such as Ag⁺ might be low on the surface of Fe-based semiconductors. On the other hand, it is noted that BiFeO₃ photocatalyst has been reported to evolve O₂ gas in the presence of Ag⁺ ions under visible light irradiation [13]. Moreover, anion-doped TiO₂ such as N–TiO₂ and S–TiO₂ are widely known to be active photocatalysts under similar irradiation for the degradation of various organic compounds, but their capabilities for O₂ evolution in the presence of Ag⁺ ions are negligible, and the photocurrents of anion-doped TiO₂ photoelectrodes are very low [22, 27].
5.3 Development of BiVO₄ Photoelectrode for Water Splitting

The fundamental principles by which photoelectrodes and photocatalysts operate are similar; however, quantum and solar energy efficiencies are much higher for the former than for the latter. As mentioned earlier, BiVO₄ semiconductors appear to be successful for use as both photoelectrodes and photocatalysts. We have found that BiVO₄ photoelectrodes are highly efficient [9], and BiVO₄ is also known to be an efficient photocatalyst for O₂ evolution from aqueous AgNO₃ [19]. The following discussion introduces the photoelectrochemical and physical properties of BiVO₄ semiconductors in detail.

First, the relationship of the photoelectrochemical properties and the morphology of the BiVO₄ film prepared by various methods was investigated.

BiVO₄ film preparation by the metal organic decomposition (MOD) method: BiVO₄ films were prepared by a modified MOD method [9]. Bismuth 2-ethyl-hexanoate in acetylacetone and vanadium (IV) (oxy)acetylacetonate in acetylacetone were mixed in a stoichiometric ratio, PEG 300 was added and the solution was concentrated by evaporation of the solvent. A spin coater was used to coat FTO conducting glass with the solution, and the coated glass was fired at 400–550°C for 30 min.

BiVO₄ film preparation by the conventional slurry (CS) method: BiVO₄ powder was prepared from K₃V₅O₁₄ and Bi(NO₃)₃ by the ion-exchange method as described elsewhere [19]. The powder was mixed with water, acetylacetone, and Triton-X in a mortar. The resulting slurry was spread with a glass rod on FTO conducting glass, and the coated glass was annealed at 500°C for 30 min.

BiVO₄ film preparation by the chemical-bath deposition (CBD) method: FTO conducting glass was soaked in a solution of Bi(NO₃)₃, NaVO₃, and EDTA at 100°C for 2 h, and the deposited BiVO₄ film was washed thoroughly with distilled water [28].

Three crystalline phases have been reported for synthetic BiVO₄: monoclinic scheelite phase, tetragonal scheelite phase, and tetragonal zircon phase [9, 19].

Which phase is present depends on the preparation method. All BiVO₄ film electrodes that we prepared were composed of the monoclinic scheelite phase, as confirmed by X-ray diffraction. Table 5.3 shows the dependence of IPCE (at 400 nm) on the preparation method and calcination temperature. The electrode prepared by the MOD method exhibits the best IPCE, and that prepared by the CS method exhibits the worst. The IPCE increases with calcination temperature up to 500°C for the MOD method, but decreases at 550°C.

Figure 5.5a–c shows scanning electron microscopy (SEM) images of BiVO₄ thin-film electrodes prepared by the MOD method at different calcination temperatures. The structure is porous, so an electrolyte solution can penetrate the film pores. The electrodes show good mechanical stability and cannot be removed by a hard scratch with a Teflon plate and a nail, suggesting good adhesion of the
film due to strong connections among particles, and between the particles and the conducting glass. The particle size increases with increasing calcination temperature, and the surface area and porosity decrease. Particle sizes at 400 and 500°C are 30–60 and 90–150 nm, respectively; the particle size at 550°C is almost the same as that at 500°C, but the porous structure of the film at 550°C is unclear. Fig. 5.5d shows a SEM image of a BiVO₄ thin-film electrode prepared by the CS method. The electrode shows poor mechanical stability and is easily removed by a soft scratch, suggesting poor adhesion among the particles, and between the particles and the conducting glass. Particle sizes range from a few hundred nanometres to 1 μm. The photocatalytic activity for O₂ evolution from aqueous AgNO₃ under visible light irradiation is higher than that of the powders used for the MOD and CBD methods. Therefore, we believe that the low IPCE obtained by the CS method is caused mainly by the difficulty of electron transfer between particles, and not by charge separation in the particles. Fig. 5.5e shows a SEM image of a BiVO₄ thin-film electrode prepared by the CBD method. The electrode is very strong mechanically and has good crystallinity, both attractive properties for charge separation and charge transfer. The particle size is very large, up to 10 μm. However, the IPCE is smaller than for electrodes prepared by the MOD method. This is attributed to the longer distance that the hole has to travel in order to reach the surface when the particle size is larger and the porosity is smaller.

Considering IPCEs and other factors, then, the MOD method with calcination at <550°C seems to be a good way to achieve porous structure and good adhesion, and hence favorable photoelectrochemical properties.

In the case of water splitting by semiconductor photoelectrodes without any sacrificial reagents in the electrolyte, measurement of O₂ and H₂ evolution is very important; however, there are very few reports on gas measurements in photoelectrochemical cells. We investigated H₂ and O₂ evolution from a BiVO₄ cell using a closed gas-circulating system. Evolution of H₂ and O₂ at a stoichiometric ratio (H₂/O₂ = 2) was observed on the BiVO₄ electrode prepared by the MOD method under visible light irradiation with the application of external bias. There was no gas evolution in the dark under the same conditions. The total amount of evolved gas was almost identical to that calculated by the total passage of photoelectrons. The BiVO₄ band structure was investigated by open-circuit potential measurement, density functional theory (DFT) calculations (Fig. 5.6), X-ray photoelectron

<table>
<thead>
<tr>
<th>Preparation method</th>
<th>Calcination temperature (°C)</th>
<th>IPCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOD</td>
<td>400</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>450</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>550</td>
<td>24</td>
</tr>
<tr>
<td>CBD</td>
<td>–</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>15</td>
</tr>
<tr>
<td>CS</td>
<td>500</td>
<td>4</td>
</tr>
</tbody>
</table>
Fig. 5.5 SEM photographs of BiVO$_4$ electrodes prepared by different methods: (a–c) MOD method at various temperatures; (d) CS method; (e) CBD method
spectroscopy (XPS, Fig. 5.7), and flat-band potential measurement [9]. The shape of the XPS valence band spectrum was similar to that predicted by DFT calculations. The BiVO₄ valence band is formed mainly by O-2p, but V-3d and Bi-6s contributions are also observed in the calculated DFT spectrum. The peak at -10 eV is mainly formed by Bi-6s, but O-2p is also observed in the DFT spectrum. It is suggested that these orbitals are hybridized with one another. Valence-band thresholds, determined by XPS for various semiconductors, are in the following order: Bi₂O₃ < BiVO₄ < (TiO₂, WO₃, V₂O₅), suggesting that the top of the valence-band potential for Bi-based oxide semiconductors is shifted negatively
compared to the potentials for conventional oxide semiconductors without Bi [9]. The speculated band structures are shown in Fig. 5.8. Hybridization between the O-2p and Bi-6s orbitals and repulsion of two orbitals, as discussed earlier and shown in Fig. 5.4, might contribute to the negative shift of the BiVO₄ valence band.

In photocatalyst studies, surface modifications and co-catalyst loading have been investigated for promotion of charge separation and formation of an active reaction site. However, reports on surface modifications of photoelectrodes are limited [6, 29–32]. An improvement in the IPCE of a TiO₂ single-crystal photocathode has been reported after adsorption of Fe³⁺ ions [32]. We investigated surface modifications of the BiVO₄ electrode by treatment with aqueous solutions (0.01 mol/L, 12 h) containing various metal nitrates and chlorides such as AgNO₃, Pd(NO₃)₂, Cr(NO₃)₃, Cu(NO₃)₂, Ni(NO₃)₂, Fe(NO₃)₃, RuCl₃, RhCl₃, H₂PtCl₆, and HAuCl₄. Among them, treatment with AgNO₃ aqueous solution improves the photocurrent significantly. The maximum IPCE for a Ag⁺-treated BiVO₄ electrode at 420 nm is 44%. This value is below only that of the nanocrystalline WO₃ electrode among oxide semiconductor electrodes under visible light irradiation and highest among mixed oxide semiconductor electrodes, suggesting great potential for porous thin-film electrodes using various mixed metal oxide semiconductors. In the case of BiVO₄ photocatalyst, high O₂ evolution activity was observed using Ag⁺ ion as an electron acceptor [19], and the degradation activity of alkylphenols was improved by loading of oxidized Ag particles (AgOₓ) as co-catalysts [33]. We investigated AgNO₃ treatment for other porous semiconductor electrodes, such as TiO₂ and WO₃, but observed no positive effects on photocurrent and stability. These results suggest a unique interaction between BiVO₄ and Ag. AgNO₃ treatment also improves photocurrent stability. We surmise that Ag⁺ ions in and on BiVO₄ catalyze intrinsic O₂ formation with the holes [9]. The observation of this positive interaction between BiVO₄ and Ag in both photocatalyst and photoelectrode is interesting.

The photocurrent and IPCE were both improved by the addition of various organic compounds such as alcohols, acids, and aldehydes (Fig. 5.9). Photocurrents were three and six times improved at 0.5 V vs. Ag/AgCl by the addition of methanol and HCOOH. The open circuit potentials were shifted negatively by the addition of
these organic compounds. The IPCE at 380 nm is >80% when HCOOH is added at 1 V vs. Ag/AgCl. The absorbed-photon-to-current conversion efficiency (APCE), calculated by light absorption loss, is >120%. Therefore, both the hole scavenger effect and the current doubling effect occur when organic compounds are added. We suggest that a photoelectrode system in combination with photodegradation of organic substances can be useful for both environmental purification and energy production.

5.4 Development of A System for High-Throughput Screening of Mixed Metal Oxide Semiconductors for Photoelectrodes and Photocatalysts

In previous study of porous photoelectrodes, we and other groups have mainly adapted mixed metal oxide semiconductors developed for photocatalysts into photoelectrodes. During this process, we have recognized some intimate relationships between photoelectrodes and photocatalysts. For example, for O₂ evolution photocatalysts with sacrificial reagents, most semiconductors are adaptable into porous photoanodes for O₂ evolution, as shown in Table 5.2. Therefore, we conclude that we should continue developing new semiconductor materials for use as both photoelectrodes and photocatalysts. However, it takes many years to manually find efficient semiconductors composed of multiple metals; therefore, we developed a system for high-throughput screening of new visible-light-responsive semiconductors for photoelectrodes and photocatalysts [34].

We selected photoelectrochemical measurements as the method for evaluating the charge-separation ability, and developed an automatic semiconductor synthesis system for preparing porous thin-film photoelectrodes of various materials. The system consists of an automated liquid-handling platform, electric furnace, and
robot arm to transfer samples, and can be used to prepare semiconductors with high element-ratio accuracy. The system uses two types of evaluation programs and two types of libraries. One program is for high-speed evaluation (Fig. 5.10): various samples are printed on the same conductive glass substrate, and their photocurrents are evaluated quickly by scanning light on the substrate. Another program is for detailed evaluation: samples are printed on separate substrates and evaluated individually. The libraries are evaluated photoelectrochemically using focused light and an X–Y stage.

An example of system screening is as follows. An n-type semiconductor, Fe₂O₃, was selected as a target material, and binary iron oxides such as Fe–Ti, Fe–Nb, and Fe–V, with various ratios were synthesized. The Fe–Ti and Fe–Nb binary oxide systems have been reported previously [6], and the result obtained from our system was consistent with those reports, demonstrating the usefulness of our system. For the previously unreported Fe–V system, the highest photocurrent was observed with 50% V, corresponding to FeVO₄, a potential new visible-light-responsive material.

Another example of system screening, among iron-based ternary oxides, is shown in Fig. 5.11. We observed excellent photocurrent for the Fe–Ti–Zr system; the highest photocurrent was observed at Fe/Ti/Zr = 60:10:30 with and without UV-cut filter. In yet another example, we screened for a p-type semiconductor material of Bi binary oxides and found CuBi₂O₄ to be a new visible-light-responsive p-type semiconductor [34, 35]. Recently, we developed a highly efficient, visible-light-responsive photocatalyst by combining the p-type semiconductor CuBi₂O₄ and the n-type semiconductor WO₃. The combined CuBi₂O₄/WO₃ shows higher reactivity than do typical TiO₂-based photocatalysts for complete oxidation of acetaldehyde into CO₂ under both solar-simulating light and visible light irradiation. The photoelectrochemical measurements of CuBi₂O₄ and WO₃ electrodes suggest that the photocatalytic reaction mechanism can be explained by the model p–n photochemical diode for reductivity/oxidizability improvement (Z-scheme mechanism) [35]. There are some reports on high-throughput screening systems where the semiconductor libraries were prepared by electrochemical deposition or ink-jet printing [36, 37]. We suggest that some new materials developed in some high-throughput screening systems using photoelectrochemical measurements can prove
useful in not only photocatalysts for water splitting but also photocatalysts for degradation of harmful compounds or for self-cleaning.

Acknowledgment  This study was partially supported by the New Energy and Industrial Technology Development Organization (NEDO) of Japan.

References


15. Abe, R., Tomita, T., Sugihara, H., Domen, K.: The use of \( \text{TiCl}_4 \) treatment to enhance the photocurrent in a \( \text{TaON} \) photocathode under visible light irradiation. Chem. Lett. 34, 1162 (2005)


20. Kudo, A., Hijii, S.: \( \text{H}_2 \) or \( \text{O}_2 \) evolution from aqueous solutions on layered oxide photocatalysts consisting of \( \text{Bi}^{3+} \) with 6s configuration and \( \text{d}^0 \) transition metal ions. Chem. Lett. 28, 1103 (1999)


Chapter 6
Combinatorial Identification and Optimization of New Oxide Semiconductors

Bruce A. Parkinson

6.1 Background

The world must eventually shift from an energy economy primarily based upon fossil fuels to one based on renewable energy sources due to the eventual depletion of fossil energy supplies and concerns about climate change due to the accumulation of carbon dioxide in the atmosphere. Sunlight is the only renewable energy source that has the potential to provide the enormous amounts of energy required by the increasing population and economic development of the world. The production and efficiency of photovoltaic solar cells that produce high value electrical energy directly from sunlight are steadily increasing. However, electrical energy is not easily or economically stored. Since 75% of the world’s present energy use is in the form of fuels, methods must be developed to transform solar energy directly into storable fuels. The photoelectrolysis of water directly with sunlight is the most promising method for achieving this result. Hydrogen, the most energy-dense fuel known, can be obtained from a readily available source and, when burned or used in a fuel cell, produces only water in a reversible cycle. Therefore, the development of a cost-effective method for splitting water directly with sunlight would pave the way to a future hydrogen economy, where no pollution or greenhouse gases are produced and that is based on a universally accessible energy source.

Current technology is capable of efficiently producing hydrogen from sunlight and water by connecting conventional solar cells to an electrolyzer [1]. The major disadvantage of this approach is that both of these devices are expensive. In addition, if the electrolyzer is not located near the solar cells, losses will be incurred during transmission of the electrical power, conversion from direct current to alternating current for transmission, and the voltage step-up and step-down at the solar array and electrolyzer,
respectively. A potentially more efficient approach is to combine these two expensive components into an inexpensive single device that both absorbs the solar energy and performs the oxidation and reduction of water. In this approach, the generated hydrogen can be transported through pipelines (with significantly less loss than the present transport of electrical power on transmission lines) or in high-pressure tanks.

Although oxygen evolution was observed as early as 1968 when illuminating a rutile electrode in solution [2], application of this concept to water photoelectrolysis was first pointed out by Fujishima and Honda in a series of experiments that used the n-type semiconductor rutile form of TiO₂ [3, 4]. Although TiO₂ is stable under illumination in aqueous electrolytes, its large bandgap (3.0 eV) limits its utilization to the UV portion of the solar spectrum and thus limits its ultimate efficiency. It should also be pointed out that the conduction band of rutile is not negative enough to reduce water and so in the original publications a “pH bias” was present where the oxygen-producing side of the cell was basic with respect to the hydrogen-producing electrolyte. Later work using heterojunctions of III–V materials as photoelectrodes considerably increased the visible light conversion efficiency but also increased the cost and decreased the effective lifetime of the device due to corrosion in aqueous electrolytes [5]. Stability of a photoelectrolysis electrode is crucial for a viable system since the significant capital investment needed for the support structures, electrolyte handling and gas handling systems requires that the system last for many years. Since the photoelectrode will be continually immersed in an electrolyte and illuminated with direct solar radiation, a long lifetime is a daunting challenge that must be adequately addressed even in the basic research stages. It is primarily for their stability that metal oxide semiconductors arguably hold the most promise for constructing a stable photoelectrolysis system.

In addition to low cost and stability, any efficient photoelectrolysis material must also effectively utilize the solar spectrum and generate sufficient photovoltage to drive the water-splitting reaction. Either a single or tandem semiconductor electrode system may be used [6]. The obtained photovoltage must be greater than the thermodynamic value for the difference between the water oxidation and water reduction potentials (1.23 eV). The material(s) should have some catalytic activity for hydrogen or oxygen production so as to minimize the additional photopotential above 1.23 V that is required to overcome the electrochemical overpotentials needed to drive the water oxidation and reduction reactions at the desired rate. Current densities (rates) on the order of 15–25 mA/cm² at illumination intensities around 1 sun are also needed for a single photoelectrode system and half these values for a tandem system. In addition, a single semiconductor photoelectrode must have a valence band edge that is more positive than the water oxidation potential and a conduction band edge that is more negative than the water reduction potential so that the photogenerated holes and electrons have the electrochemical driving force for the water-splitting reaction. If a dual photoelectrode system is used, the conduction band of the p-type material must be negative of the water reduction potential and the valence band of the n-type material must be positive of the water oxidation potential.

As stated above, either a single semiconductor electrode or a two-semiconductor electrode photoelectrolysis system can be envisioned. The single illuminated
electrode system will need a bandgap greater than about 1.7 eV in order to supply the photovoltage needed for water photoelectrolysis given that photovoltage of about 2/3 the band gap is often the maximum obtainable from a good semiconductor material operating at its maximum power point. A single photoelectrode system would need to be attached to a catalytic metal (perhaps Pt or Pt modified) electrode that can accomplish the complementary water-splitting half reaction. A two-semiconductor system, with both a p-type and an n-type semiconductor photoelectrode, can be configured in several ways. In one configuration the electrodes could be placed with the larger gap material absorbing the higher energy portion of the solar spectrum in front of the smaller bandgap material of the opposite carrier type in which the wavelengths transmitted through the large bandgap material would be absorbed in the smaller gap material. The two materials must absorb nearly equal numbers of solar photons since the current through each semiconductor must be matched due to the recombination at the ohmic contacts that provides the summation of the two photovoltages. Therefore, in this configuration the quantum yield for water photoelectrolysis will be half that of a single photoelectrode system; however, since there is considerable photon flux in the red and near IR portion of the solar spectrum, the lower quantum yield may be offset by utilizing this portion of the solar spectrum. This configuration has the advantage that one glass substrate, with a transparent conducting layer on both sides, could be used for both materials reducing the total system costs. The ideal bandgaps for the two materials have not yet been identified but they will most likely be somewhat different than for a two-bandgap heterojunction solar cell where bandgaps of 1.9 and 1.2 eV give the best theoretical light to electricity conversion efficiency but yield a photovoltage larger than necessary for water photoelectrolysis. A larger photovoltage would provide some extra driving force for a photoelectrode that does have optimal catalytic properties for water oxidation or reduction. A schematic of a cross section of such a device is shown in Fig. 6.1. Systems that convert the extra photovoltage to electrical power can also be designed but the electrical power produced would most likely be more expensive than if produced from a conventional photovoltaic device.

Another configuration would be to place the n and p type materials side by side rather than putting a larger bandgap material in front resulting in full spectrum sunlight illumination for both electrodes. In this case, an n-type and p-type electrode of the same material or two different oppositely doped smaller bandgap materials can be paired as long as the sum of their photovoltages at maximum power was greater than the ~1.7 eV needed for water photoelectrolysis and the photocurrents in both electrodes were matched. If two materials were used the bandgaps would have to be close to each other since they need to absorb nearly the same flux of solar photons in order for the photocurrents to be matched. Two stable materials would be needed for this configuration to work, a p-type material with a conduction band at least 0.8 eV negative of both the n-type material conduction band and the hydrogen potential and valence band offsets of about 0.8 V. Two substrates would also be required, but they need not be transparent (as in the case of the stacked system discussed above and shown in Fig. 6.1) and so
cheaper and more abundant materials may be utilized. Table 6.1 summarizes the advantages and disadvantages of the different photoelectrolysis configurations discussed above. Systems that contain buried junctions, where charge separation is remote from the electrolyte, are really solid state solar cells in series with metal electrodes, are not considered since they are essentially equivalent to hooking up photovoltaic cells to an electrolysis cell.

Any photoelectrolysis system will have to be designed with an ion conducting separator or electrolyte path between the two electrodes and with separate gas collection plumbing for hydrogen and perhaps oxygen. Oxidation reactions, other than water oxidation, have been discussed since there is much less demand for pure oxygen than there is for hydrogen. The commercial demand for pure oxygen would be rapidly saturated once any water-splitting photoelectrolysis system with the capacity to produce meaningful amounts of hydrogen is deployed. The only other commercial chemical currently produced on a very large scale by aqueous electrolysis is chlorine, but again the demand for chlorine would be quickly met by any system producing enough hydrogen fuel to make an impact on world energy demand. However, photoelectrolysis technology could get an initial boost from the additional revenue obtained from the photoelectrolysis of hydrochloric acid or brine.
Often materials are synthesized and tested for photochemical water-splitting activity as either colloidal solutions or as powdered slurries. While this method may be useful for screening for water-splitting activity, it is unlikely that homogeneous colloidal solutions or slurries will be useful for a practical water-splitting system. Several reasons for this include (1) The products are not produced in separate compartments, resulting in highly explosive mixtures of hydrogen and oxygen. (2) Energy is required to separate hydrogen and oxygen reducing the overall efficiency of the water-splitting process. (3) Since catalysts for hydrogen and oxygen production from water are also good catalysts for the recombination of hydrogen and oxygen, illumination of such systems will eventually result in a photostationary state where forward and back reactions have equal rates and no more water splitting can occur. The exchange current at the photostationary state will be related to the solubility of the gases in the electrolyte and the rate at which they are removed from the reactor.

To avoid undesirable back reactions, and the poor kinetics for hydrogen or oxygen production, sacrificial reagents are often used in the photoreactors to scavenge either holes or electrons. The most common examples of hole scavenging agents introduced into solution are alcohols (usually methanol), amines (usually triethanolamine or EDTA), or sulfite salts. Electron scavengers such as the easily reduced Ag⁺ can also be added. Electron or hole scavengers can be useful to study one of the water-splitting half-reactions without complications associated with the kinetics of the other half-reaction (although using a three-electrode potentiostat to study a single photoelectrode would achieve the same end) but these additives are not viable for any practical system for sustainable energy production.

As a comprehensive review of photoelectrochemical hydrogen generation [7] and a recent book [8] have been published on this subject, this chapter will concentrate on the combinatorial aspects of the search for appropriate materials for a practical device.

### Table 6.1  A comparison of the advantages and disadvantages of various configurations for a photoelectrolysis device

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single photoelectrode</td>
<td>Only one substrate and one p or n type semiconductor needed. Dark electrode can be conventional.</td>
<td>Poor utilization of the solar spectrum since large bandgap (≥1.7 eV) is needed.</td>
</tr>
<tr>
<td>Tandem photoelectrodes (Fig. 6.1)</td>
<td>Needs only one substrate. More efficient utilization of the solar spectrum.</td>
<td>Need to identify one p and one n type semiconductor and needs a transparent conducting substrate. Two photons and current balancing needed.</td>
</tr>
<tr>
<td>Separately illuminated photoelectrodes</td>
<td>Cheap nontransparent substrates can be used. More efficient utilization of the solar spectrum.</td>
<td>Need two semiconductors or ability to dope one both n and p type. Two photons and current balancing needed. Twice the collector area of tandem.</td>
</tr>
</tbody>
</table>
6.2 Survey of Known Materials and Their Photoactivity

As discussed above, the positions of the conduction band and valence bands with respect to water oxidation and reduction potentials are one requirement for a water-splitting photoelectrode. Among the first materials to be exhaustively studied were the perovskite structure oxides SrTiO$_3$ [9–13] and BaTiO$_3$ [9, 14, 15]. Both of these materials are better suited than TiO$_2$ for water photoelectrolysis because their conduction band edges are more negative than the water reduction potential. Unfortunately, as is also true for TiO$_2$, their bandgaps are so large (~3.2 eV) that they also required UV light to produce photocurrent resulting in very poor solar efficiencies.

A practical solar photoelectrolysis material must also be stable under illumination for many years while immersed in an aqueous electrolyte. Photocorrosion is the main problem with nonoxide semiconductors that have near optimum band gaps [16] and band positions for efficient photocatalytic water splitting such as transition metal chalcogenides [17–19], cadmium and zinc chalcogenides, GaAs [20], GaP [21], and InP [5, 22–24]. Thermodynamic stability is assured only if the oxidation decomposition potential of the semiconductor is more positive than the valence band edge and the reduction decomposition potential lies more negative than the conduction band edge [25]. Because the redox potentials for the decomposition reactions of these materials are generally within the band gap, and additionally because the decomposition reactions are kinetically favored, these experiments generally lead to electrode photocorrosion. If even only a tiny fraction of the photoexcited carriers are involved in photocorrosion reactions the electrode will not survive for many years under solar illumination intensities.

The most stable rocks and minerals that survive many years of weathering are all oxides, suggesting that a semiconducting oxide material will have the best chance to fulfill the stability criteria for a water-splitting photocatalyst. Smaller bandgap metal oxide semiconductors, notably WO$_3$ [26–29], Fe$_2$O$_3$ [30–35], NiO [36, 37], and Cu$_2$O [38], have also been extensively studied. Although these materials harvest a larger percentage of the solar spectrum, their conduction band edge energies are often not optimally positioned for the production of hydrogen [39]. They can also be difficult to electrically dope, resulting in being either poor conductors of electrons and/or holes and some of them (e.g., Cu$_2$O) are still prone to corrosion [16].

The valence band positions of oxide materials are usually similar and quite positive due to the fact that the band is derived from oxygen 2p states. Holes at such a positive potential can easily oxidize water but can waste considerable energy in doing so either due to the large difference in potential between the holes and the water oxidation potential or poor electrochemical kinetics for water oxidation. One idea to both extend the spectral response of large gap oxide semiconductors and to raise the valence band energy is to incorporate sulfur, carbon, or nitrogen into the lattice. A series of these compounds including Sm$_2$Ti$_2$S$_2$O$_5$ [40], TaON [41], Y$_2$Ta$_2$O$_5$N$_2$ [42], Ta$_3$N$_5$ [43], and LaTiO$_2$N [44, 45] as well as incorporating S, N, or C into TiO$_2$ [46–49] have recently been examined by several research groups. These groups often use density functional theory to calculate that the bottom of the
Conduction bands consist of empty d-orbitals and, depending upon the material, the tops of the valence bands consist mainly of N 2p or S 3p orbitals resulting in smaller band gaps for improved solar energy utilization. The smaller band gaps, and fortuitous band positions relative to the redox potentials of water, imply that these materials could function as water photoelectrolysis materials; and indeed H₂ and/or O₂ is produced at quantum efficiencies up to 30% [41]. Unfortunately, however, most of these systems are typically not reported to function without the presence of the “sacrificial reagents” in solution mentioned earlier. It is likely that the sacrificial reagents are necessary to camouflage inherent oxidation or reduction decomposition reactions and that these materials will ultimately have the same stability problems as the earlier materials that incorporated nonoxide anions since nitrogen, sulfur and carbon are easily oxidized in solution.

In principle, if one introduces an additional metal or metals into the structure that have electronegativities greater than that of the parent metal and less than that of oxygen, then it is possible that a valence band created from mixed metal d states will be more negative in energy (closer to the oxidation potential of water) than that of a pure oxygen 2p band, thereby reducing the bandgap to 1.5–2.0 eV [50]. As an example, a mixed metal oxide semiconductor could be constructed with conduction band edge derived mainly from titanium. In principle, it should then be possible to create mixed bands, lowering the band gap of the parent titanium oxide, by mixing the transition metal orbitals with the oxygen 2p orbitals. For example, the ions Mn²⁺, Ni²⁺, Cr³⁺, Rh³⁺, V⁴⁺, Os³⁺, Sn²⁺, Sb³⁺, Pd²⁺, and Pt²⁺ all have d orbital energies that are more positive (further down an absolute scale) than the oxidation potential of water but more cathodic than that of oxygen’s 2p orbital [50]. And so, in theory at least, by combining these metals with titanium it should be possible to create a material with a band gap that can absorb a large fraction of the solar spectrum and is optimally positioned to photoelectrolyze water. Other compositions such as LuRhO₃ [51], SrZr₀.₂₅Ti₀.₇₅O₃ [50], LaVO₃ [52], K₂La₂Ti₃O₁₀ [53], and RbPb₂Nb₃O₁₀ [54], have also been reported to be viable photoelectrolytic systems and provide useful starting points in a combinatorial search.

6.3 Combinatorial Production and Screening of Metal Oxide Photocatalysts

Multicomponent materials are often required to optimize a special property. High temperature superconductivity is a good example of a special property requiring multicomponent materials. The oxide with the highest known transition temperature contains four or even five metals (HgBa₂CaCu₂O₆+d, Hg₀.₈Tl₀.₂Ba₂Ca₂Cu₃O₈.₃₃, Tc ~ 125 K). Since our present theoretical knowledge is insufficient to a priori calculate the behavior of such complex systems, or estimate the identity and atomic ratios of the specific metals needed to perform the efficient photoelectrolysis of water, empirical methods will be necessary to identify the best material. Thousands, or perhaps hundreds of thousands, of mixed metal oxide photocatalyst compositions...
may need to be produced and tested for efficiency and corrosion resistance until an effective material is discovered. Fortunately, the recent development of combinatorial methods has provided tools for the production of large libraries of compounds and the methods for screening these libraries for the presence or optimization of a property of interest. In this section, we will review the combinatorial approaches that have been applied to the discovery and optimization of water-splitting photocatalysts.

We can first do an analysis of the number of possible combinations of mixed metal oxides that could be produced and screened. If we prepare all possible ternary oxides, mixing \( N \) materials three at a time there are \( N!/3!(N - 3)! \) combinations if we only consider 1:1:1 stoichiometries. Since there are \( \sim 60 \) possible candidate metals in the periodic table this results in \( >34,000 \) combinations to be produced and tested. If a quaternary material is needed then there are \( N!/4!(N - 4)! > 480,000 \) possible combinations. Again these numbers do not consider producing and screening the many possible stoichiometries. It is also possible that a ternary or quaternary material will require small amounts of an additional element that acts as an electrical dopant to optimize its conductivity. These small amounts can often exist as impurities in the metal oxide precursors or may have to be discovered by additional combinatorial screening (discussed later in this chapter) greatly multiplying the number of experiments that must be done. However, some combinations could be excluded because it would not make sense to, for example, mix only large bandgap oxide materials (i.e., Ti–Zr–Si–Y oxides) and expect visible light absorption. It is clear that given the large number of possible mixtures that need to be produced, a “high-throughput” approach is needed, as well as multidimensional experimental design, to explore the parameter space more rapidly. And while it may not be reasonable to produce and test ALL of the possible combinations, it is certainly imperative that the throughput of materials screening be rapid relative to the serial approaches that have historically been used.

We can also use some chemical knowledge to reduce the number of combinations that need to be tested. We can very loosely classify metals into one of four categories according to their potential role in a photoelectrode material: structural, since they form very stable high heat of formation oxides (i.e., Ti, W, Zr, Ta, Si, Mo, Nb, Hf, In, Sn, Y, Sc whose oxides have very high bandgaps and so are not useful by themselves), light absorbing, where multiple elements may be needed to absorb light across the solar spectrum [55, 56] (i.e., Fe, Cr, V, Co, Mn, Ni, Cu, and some rare earths such as Ce), catalytic, to lower overpotential losses in the electrochemical reactions (i.e., Ru, Rh, Pd, Ni, Pt, Ir, Os, Re) [55, 57, 58] and ionic charge compensators (i.e., Ca, Sr, Ba, Mg, Zn, Cd, Li, Na, K, Rb, Cs)\(^1\) that are often present in stable structures such as spinels and perovskites to balance the charge in the compound since oxygen is almost always present in stable oxides as \( O^{2-} \).

---

1 We purposely do not include toxic metals such as Pb, Tl, Cd, and Hg since we envision eventual large-scale implementation of any discovered photocatalysts and we want to avoid the environmental consequences of these elements.
This chemical guidance greatly reduces the number of possible combinations (but it is still in the tens of thousands). Finally, it may be necessary to deposit catalytic materials such as Pt, Pd, Ir, Ru, or Ni [59–61] or small molecule catalysts onto the surface of the electrode if the surface of the semiconductor itself has little catalytic activity.

With the appropriate substrate a combinatorial library can be tested for photoactivity by a number of different methods. While the ideal method would be to directly measure hydrogen production with, for example, gas chromatography, the amount of hydrogen expected to be produced from a thin film photoelectrode that does not absorb a large fraction of the incident light (the reason for very thin films is discussed below) is typically too small to generate hydrogen at a rate needed to overcome the solubility of the gas in aqueous solution. In light of this limitation, screening for photocatalytic activity has largely been carried out by measuring spatially resolved photocurrents or by observation of pH changes in the electrolyte near the array.

The use of pH-dependent fluorescent indicators [62] or molecular sensitizers [63] is based upon the principle that the local concentration of protons increases at sites catalytically active for water oxidation (oxygen production) and decreases at sites active for water reduction (hydrogen production). The reverse processes of hydrogen oxidation or oxygen reduction also cause a change in pH making the method useful for screening fuel cell catalysts as well as photoelectrolysis catalysts. A visual or CCD camera inspection of an irradiated catalyst library can be sufficient to ascertain activity if the amount of redox chemistry and/or the fluorescence of the indicator is sufficient. Otherwise, a spatial analysis of absorbance or emission changes must be carried out. The advantage of this approach is that it is a rapid detection method and avoids some of the spurious signals that can result from electrochemical detection methods. On the other hand, it is not high resolution and one must assume that the indicator molecules do not interfere with the overall redox process occurring in solution (such as competitive charge transfer to the indicator in lieu of photocatalysis) and the appropriate indicator must be found for the chosen electrolyte. Alternative screening techniques with higher resolution will be described in the following section.

6.4 Methods for Combinatorial Production and Screening

6.4.1 Inkjet Printing and Electrochemical Scanning

Light Beam Analysis

Inkjet printing has been utilized for many combinatorial chemistry applications including the production of luminescent metal oxide libraries [64]. The hallmarks of inkjet printing for materials deposition are its speed, ease of use, low cost (depending upon the printing platform chosen) and its versatility. The combinatorial search strategy within our group [58], and later adopted by the Lewis group [65],
utilizes inkjet printing of overlapping patterns of soluble metal oxide precursors onto a conductive substrate. Metal nitrate salts have been the most widely printed precursors since they are highly soluble and subsequent pyrolysis at temperatures below 550°C yields metal oxides. Temperatures must be kept below about 550°C since above these temperatures the glass begins to soften and the conducting FTO layer becomes unstable.

Our work has used a research grade printer manufactured by FujiFilm Dimatix [58] and off-the-shelf consumer grade printers made by Hewlett Packard [66]. Originally, films were deposited in overlapping triangular gradient patterns as shown in Fig. 6.2. While inkjet printing such a pattern ideally produces smooth gradients in composition, in reality the extent of uneven mixing of compositions within the triangles can be unknown. Assuming that the actual deposition amounts of the material precursors is linear from the vertices to the edges of the triangle, other factors such as miscibility, phase segregation, annealing effects, etc., might also play a role and make it difficult to know the exact composition in a particular spot. One particular complication that has emerged during our studies is that the metal nitrate salts generally melt before they decompose and coalesce into droplets, resulting in less uniform films of the oxide after pyrolysis. The use of metal oxide nanoparticles, polyoxometalates, and various alkoxide compounds as precursors that decompose directly into the oxide will be useful to avoid this complication.

To analyze for photoelectrolysis activity several groups use custom-built scanning laser screening systems where a laser is rastered over the printed and pyrolyzed substrates that are immersed in an aqueous electrolyte. Scanning the
laser over the sample produces a false color image of photoactivity by measuring the photocurrent at each “pixel.” To detect small signals the laser can be chopped or modulated and a lock-in amplifier can be used to isolate the photocurrent signal even if there is substantial dark current present. A potentiostat can also be used to apply a positive or negative external bias to the sample to aid the oxidation or reduction of water, respectively. The external bias will not be useful for a practical system but is used so that a good semiconducting material that might not be electrocatalytically active may still be identified and can have a surface catalytic layer added later.

Figure 6.3 shows a template, a photograph, a false color photocurrent image and a false color photocurrent image of a compositional zoom of a promising compositional region. In the false color photocurrent images in Fig. 6.3, the yellow to white pixels corresponds to higher relative photocurrent where a Fe–Cs–Nd composition is most photoactive. The dark blue regions represent areas where there is virtually no photoactivity either from the FTO substrate or the material. Also printed and screened using this template are n- and p-type “internal standards” of α-Fe$_2$O$_3$ and CuO, respectively, printed because it not only affords a direct comparison to these well-known semiconducting oxide materials [20], but also between all printed sample plates.

Hundreds of materials have been produced and screened using the unfolded tetrahedron template, with most of the material combinations showing no or very little photoresponse but some showing significantly better performance than the internal reference materials. One of the materials discovered in our search, composed of Fe, Co, and Al, has reproducibly shown p-type photoactivity and was not previously known to function as a photoelectrolysis material. The regions of high activity were found in the Co-rich regions of the Co–Al–Fe triangles. While this original high-throughput method was useful for identifying promising materials, the exact stoichiometry of the metal components in the active mixture was uncertain since the exact amounts of the various “inks” delivered by the printer driver software is unknown.

In order to determine the optimum stoichiometry of the promising mixture, a series of samples with an array of areas with known compositions was prepared using a research grade inkjet printer manufactured by FujiFilm Dimatix. This printer has a camera that focuses on the printhead nozzles and software control of parameters such as the waveform applied to the piezoelectric jetting nozzle, drop spacing, drop volume, and jetting frequency. By controlling these parameters, production of a broad or narrow compositional range with known stoichiometries within each printed area is possible. The thickness of each sample can also be controlled with these factors or, if desired, by depositing multiple layers. An example of a sample plate prepared in this way is shown in Fig. 6.4. In this pattern, four different densities of Co and Fe were printed as double horizontal and double vertical rows, respectively, while different amounts of Al were deposited in a repeating pattern within groups of four squares containing identical Fe and Co amounts as shown in the dashed inset in the upper left-hand corner. The precise values of the drop spacing and density printed within each square in
Fig. 6.3 Printing and screening a four-metals-three-at-a-time pattern and a compositional zoom for a Fe–Cs–Nd–Cu system. (a) False color template showing the positions and gradients used for printing the four metal precursor solutions. (b) Photograph of the printed and fired film. Note the triangular internal standards of a-Fe₂O₃ and CuO (upper right and left, respectively) with thickness gradients (bottom to top) that are used as internal standards. (c) False color photocurrent image of the film shown in B using 514.5 nm illumination under 0.5 V bias in a 0.5 M NaOH solution. The photocurrent produced at a particular “pixel,” relative to the others in the two-dimensional array, is represented by its relative brightness with the most photocurrent in a particular direction (water oxidation in this case) being the brightest. (d) Photocurrent scan at 514.5-nm of a triangular composition zoom in on the brightest area of the Fe–Cs–Nd triangle shown in (c) that has a maximum IPCE value approximately twice that of the a-Fe₂O₃ internal standard (smaller triangle to the lower right). Expanding the printing gradients within the brightest region of the n-type material created the “zoom in.” Reproduced with permission from [66], copyright 2005 American Chemical Society

Fig. 6.4 for the Fe and Cu internal standards, as well as the Fe–Co–Al Oxide precursors, can be found in reference [66]. As an example, the brightest square in Fig. 6.4, third square from the right, third row down, is given the label Fe₇Co₁₄.₄Al₁ (In atomic percentages 10% Fe, 84% Co, and 6% Al) and this composition was used in future photoelectrochemical and structural
experiments where bulk samples were prepared as it consistently appeared to be a near optimal composition.

In contrast to the “zoom in” shown in Fig. 6.3, created by adjusting the printed gradients, the most promising stoichiometries within array-based samples can be investigated much more quantitatively. To further refine the optimal composition of the photocatalyst described above, another template was created that encompasses stoichiometries outlined by the white triangle as well as some similar stoichiometries. The template, and the result of the photocurrent screening, is shown in Fig. 6.5. In this pattern, the amount of cobalt was kept constant (labeled B in the figure with a drop spacing equal to that of all mixtures within the white box of Fig. 6.4) while the Fe and Al contents (labeled w, x, y, and z) were varied and calculated using the rationale described previously. The internal standards were printed with the same drop densities as in Fig. 6.4 and the overall amount of printed internal standard material relative to the working mixtures is approximately the same between Figs. 6.4 and 6.5 in order to not overestimate the contribution of sample thickness to performance. The resulting photocurrent for all working Co–Al–Fe oxides is within $\pm 20\%$ for all squares indicating flexibility in the stoichiometries of the photoactive compositions. Indeed, the Fe and Al content can vary by as much as a factor of two (between Co$_{17.8}$Fe$_{1}$Al$_{2.1}$ and Co$_{17.8}$Fe$_{2.1}$Al$_{1}$) and still produce a material with comparable photoactivity.

Currently inkjet printing allows one to rapidly produce and screen many materials and quantify the most promising stoichiometries. Lewis, Katz and Gingrich [65] have devised a variation of the technique described above that allows
for the measurement of the open-circuit photovoltage as well as the photocurrent response of each composition. The photovoltage provides a valuable estimate of the position of the quasi-Fermi level of the majority carriers under illumination, which indicates whether the photogenerated carriers are thermodynamically capable of a particular photoelectrolysis reaction. Since the photovoltage measurement is for the entire electrically connected substrate, to measure it for a specific area or composition requires electronic isolation of each metal oxide mixture, as well as passivation of the exposed conductive substrate. To achieve this up to 260 regions on the FTO substrate are laser etched to remove the FTO from the regions between the areas to be printed and electrically isolate them. Up to eight precursors are then printed in the isolated regions and pyrolyzed. An insulating epoxy is then screen printed over any exposed FTO. A multiple contact connector is then clamped to the substrate and connected to a custom-designed printed circuit board where each contact on the circuit board is associated with a particular spot in the materials library and connected through ribbon cables to a multiplexer for the electrical measurements.

The Sayama group also currently utilizes a scanning light beam analysis system for high-throughput screening of photoelectrolysis materials [67]. In their system the libraries are produced by an automated robotic pipetting of precursors onto a substrate with a subsequent firing. The photocurrent of each sample in an array is then measured at a constant potential in a three-electrode cell by scanning the photoelectrode with a focused beam from a xenon lamp guided by an optical fiber with a UV cutoff filter. Their high-throughput production and screening system has been used to optimize the stoichiometry of some porous iron oxide binary compositions.

**Fig. 6.5** Compositional library produced by expansion of the compositional region outlined by the white box in Fig. 6.4. The template shown in (a) was created by including all of the mixtures created within the inset of Fig. 6.4 as well as some intermediate stoichiometries. The pure Fe and Cu internal standards were printed as in Fig. 6.4. (b) Photocurrent map of the array of 3 × 3 mm² measured in 0.1-M NaOH with a 532-nm laser showing that a rather large range of composition space appears to work better than the CuO internal standards. Reproduced with permission from [58], copyright 2008 American Chemical Society.
6.4.2 Combinatorial Dopant Optimization Using Inkjet Printing

Low levels of impurities are well-known to control the electrical properties of semiconductors via electrical doping or to produce potent deep trap states that can promote carrier recombination and destroy the photoresponse of the material. It is often difficult to control or determine these low levels of active impurities; however, in some photoelectrode materials purposely adding impurities can dramatically improve the photoresponse. One such material, hematite (α-Fe₂O₃), has received much attention because of its bandgap of 2.2 eV, abundance, nontoxicity, and electrochemical stability. Many methods have been investigated to improve the photoelectrochemical performance of α-Fe₂O₃ including the incorporation of various other elements into the hematite films. Grätzel found that the introduction of a small amount of Si into the film dramatically increased the photooxidation current. Many other “dopant” species have also been added including Ca²⁺, Mg²⁺, Cu²⁺, Zn²⁺, Si⁴⁺, Ge⁴⁺, Ti⁴⁺, Pt⁴⁺, Cr⁴⁺, V⁵⁺, Nb⁵⁺, and Mo⁶⁺ [68–82]. A number of these elements have improved the photoresponse of α-Fe₂O₃ but the mechanism for the enhancement has not often been explained. Some reasons for improvements of the photoactivity from adding impurities include lowering the resistivity of the film due to an increase in the majority carrier density, improvements in the electrocatalytic activity for water (OH⁻/CO) oxidation, directing the morphology of the film (nanostructuring) to allow more carriers to reach the semiconductor/electrolyte interface, passivating grain boundaries or producing different crystallographic orientations of the α-Fe₂O₃ crystallites to favor the growth of faces with a lower surface recombination velocity. The use of the term “dopants,” when referring to the addition of more than parts per thousand levels of impurities, does not conform to the usual semiconductor physics definition of electrically active shallow dopants that can control the electrical conductivity at the level of parts per million. However, since this term has come into wide use, even when adding amounts of elements at levels more properly called solid solutions (if in fact all the additive is incorporated into the lattice), we will continue to use the term “doping” with the understanding that the reason these added elements influence the photocurrent and/or photovoltage are not usually known.

Given the multitude of possible enhancement effects, adding multiple elemental impurities might result in synergistic effects and produce films with much higher photoactivity than with just one impurity. However, given now the large number of possible additives at multiple concentration levels the number of samples that must be prepared becomes very large. Therefore, we adapted our inkjetting technique to quickly produce many “doped” α-Fe₂O₃ compositions and screen them for their water photooxidation activity. Since we will be printing arrays of α-Fe₂O₃ we first optimized the thickness of printed films of this material by varying the drop density and the number of times we overprint the pattern. An 8 × 8 array of squares was printed onto SnO₂:F-coated glass with 0.5 M FeCl₃ with drop spacing of 80 μm. This pattern, shown in Fig. 6.6a, was divided into four double rows printed once, twice, three times, and four times from the bottom to the top, to determine the
Fig. 6.6  (a) Printing template used for undoped α-Fe₂O₃ with ~80 μm drop spacing and double rows printed from the bottom up: once, twice, three times, and four times corresponding to Feₐ to Feₜ.  (b) Printing template used for the Fe–M \( (M = \text{Ti, Si, or Al}) \) series where known amounts of the components are printed on top of the Fe rows from (a) into individual \( 3 \times 3 \) mm² resulting in groups of four squares with identical compositions. The density of precursor drops and resulting stoichiometries for (d–f) are listed in Table 6.2. False color photocurrent maps of libraries printed using template (b): (d) Ti-doping; (e) Si-doping; (f) Al-doping. All libraries were scanned at 532 nm with +0.6 bias in a 0.1-M NaOH solution. Note that the false colors are relative within a scan and not absolute values therefore comparisons of color between scans are not quantitative. Refer to Table 6.3 for more quantitative comparisons.
optimum $\alpha$-Fe$_2$O$_3$ thickness for the maximum photocurrent signal. It was found that
printing this concentration of FeCl$_3$ three times gave the best photocurrent signal
(Fig. 6.6c). The top double row (FeD) was too thick resulting in darker areas in the
center of the squares and bright edges due to the film being slightly thinner near the
edges. The short diffusion length of carriers in this material results in decreases in
photocurrent when the film is too thick. Note that in all cases most of the light still
passes through the very thin film of $\alpha$-Fe$_2$O$_3$ and so incident photon current
efficiency (IPCE) values are low.

According to the literature [68, 72, 83], the addition of small amounts of Ti, Si or
Al improve the photoelectrochemical performance of $\alpha$-Fe$_2$O$_3$. Therefore, we
selected Ti, Si, and Al to investigate the effects of the incorporation of small
amounts of these elements and combinations of them on thin film $\alpha$-Fe$_2$O$_3$
photoanodes. Combinatorial printing templates such as those shown in Fig. 6.4
were created where the total amount of precursor deposited into each square is
determined from the concentration of precursor salts in the printing solution, the
drop volume of a printed drop, the number of layers printed and the number of drops
inkjetted per unit area. This amount is determined by the areal density of dots and
the number of ink-jetting nozzles actually firing on each printer cartridge. In the
pattern, in Fig. 6.6b, four different densities of Fe and other metal (M = Ti, Si, or
Al) were printed as double horizontal and double vertical rows, respectively,
resulting in groups of four squares of nominally identical composition. The result
of the photocurrent screening is shown in Fig. 6.6d–f.

Controlling the density of the printed drops allows the production of a quantifiable
broad or narrow compositional range of dopant additions on each substrate. Table 6.2
contains the values of the drop spacings and densities used in Fig. 6.6d–f for the Fe
and M (M = Ti, Si, or Al) oxide precursors taken individually. For example, the
bottom four squares on the far left, corresponding to Fe$_{A}$M$_{D}$, were printed with a drop
spacing of 160 $\mu$m for Fe (three layers) and 40 $\mu$m for M (M = Ti, Si, or Al for
Fig. 6.6d, e and f, respectively). This set of squares corresponds to a nominal
stoichiometry of Fe$_{1}$M$_{0.095}$ or 8.7% dopant, whereas the smallest percentage square
corresponds to Fe$_{D}$M$_{A}$ with a nominal stoichiometry of Fe$_{1}$M$_{0.00047}$ or 0.047%
dopant. The corresponding numerical maximum IPCE data are summarized in
Table 6.3. These values can be used to evaluate the relative responses when compar-
ing the false color photocurrent scans since the brightest yellow shade was used to
represent the maximum photocurrent response in each image. We found that
in general only Ti-incorporation enhanced the photoelectrochemical performance

<table>
<thead>
<tr>
<th>Row label</th>
<th>ID DS ($\mu$m)</th>
<th>Drops per mm$^2$</th>
<th>Column D</th>
<th>Column C</th>
<th>Column B</th>
<th>Column A</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>40</td>
<td>642</td>
<td>0.0067</td>
<td>0.0017</td>
<td>0.00078</td>
<td>0.00047</td>
</tr>
<tr>
<td>C</td>
<td>79</td>
<td>169</td>
<td>0.025</td>
<td>0.0067</td>
<td>0.003</td>
<td>0.0018</td>
</tr>
<tr>
<td>B</td>
<td>120</td>
<td>75</td>
<td>0.057</td>
<td>0.015</td>
<td>0.0067</td>
<td>0.004</td>
</tr>
<tr>
<td>A</td>
<td>160</td>
<td>45</td>
<td>0.095</td>
<td>0.025</td>
<td>0.011</td>
<td>0.0067</td>
</tr>
</tbody>
</table>

$DS$ drop spacing
in comparison with an undoped iron oxide thin film, while the incorporation of Si or Al shows somewhat detrimental effects although there seems to be an increase in response for the thinner iron oxide films with the smallest amount of Si incorporation (FeASiA or Fe1Si0.0067 - bottom right corner of Fig. 6.6e) suggesting that perhaps even smaller additions of Si might improve the relative photoresponse. In previous work for films prepared by ultrasonic spray pyrolysis the corresponding nominal stoichiometry showing highly improved photoresponse is Fe1Si0.01 or ~1% dopant [37]. As for the Ti-doped iron oxide thin films, the best photocurrent response was observed at the relative stoichiometry of Fe1Ti0.025 while it ranges from Fe1Si0.0018 (or 0.18% dopant) to Fe1Si0.025 (or 2.4% dopant) for Si-doping, and Fe1Al0.0067 for Al-doping.

There have been numerous reports about the effects on photoelectrochemical response of incorporation of small amounts of the other elements into the binary \(\alpha\)-Fe2O3 system. However, it is challenging to investigate the effects of adding multiple elements to \(\alpha\)-Fe2O3 thin films especially if individual films have to be prepared at each stoichiometry. Our high-throughput combinatorial method can be used to quickly measure the effects of additions of multiple elements to find an optimal composition. Therefore we printed films with combinations of Ti, Si, and Al simultaneously added to \(\alpha\)-Fe2O3. A printing template shown in Fig. 6.7 was created where four different densities of Ti and Si were printed as double horizontal and double vertical rows, respectively, and different amounts of Al were deposited in a repeating pattern within groups of four squares containing identical Fe, Ti, and Si amounts. The template was printed four times with different iron precursor drop spacings. The result is that the total thickness of the \(\alpha\)-Fe2O3 decreases from Fig. 6.8a–d, whereas the relative concentration of the added dopants increases as we make the \(\alpha\)-Fe2O3 film thinner. The top row of squares in the template are equal thickness squares of \(\alpha\)-Fe2O3 where the first three pairs of squares have only one added metal dopant at the A and D levels (Al, Si, Ti left to right) with the far right pair of squares being just \(\alpha\)-Fe2O3 printed at 40 and 160 \(\mu\)m spacings (FeD and FeA). Table 6.4 contains the values of drop spacings and drop densities used to print the template in Fig. 6.7 for the Fe–Ti–Si–Al oxide precursors. For illustrative purpose, the bottom square on the far left of the template, labeled SiDTiAAlB, was printed with a drop spacing of 40, 160, and 120 \(\mu\)m for Si, Ti, and Al, respectively, while the Fe precursor was printed in three layers with different drop spacings of 40, 80, 120, 160 \(\mu\)m for Fig. 6.8a–d, respectively. So this lower left square in Fig. 6.8a–d will correspond to stoichiometries of Fe1Si0.0067–Ti0.00047Al0.00078, Fe1Si0.025Ti0.0018Al0.003, Fe1Si0.057Ti0.0049Al0.0067 and Fe1Si0.095Ti0.0067Al0.011, respectively. This printing pattern results in libraries of decreasing \(\alpha\)-Fe2O3 thickness but increasing doping levels from Fig. 6.8a–d. The results of the photocurrent screening of these libraries are shown in Fig. 6.8a–d.

<table>
<thead>
<tr>
<th>Undoped Fig. 6.6c</th>
<th>Ti-doped Fig. 6.6d</th>
<th>Si-doped Fig. 6.6e</th>
<th>Al-doped Fig. 6.6f</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 V bias</td>
<td>2.60 \times 10^{-4}</td>
<td>4.06 \times 10^{-4}</td>
<td>2.58 \times 10^{-4}</td>
</tr>
<tr>
<td>+0.6 V bias</td>
<td>11.47 \times 10^{-4}</td>
<td>15.54 \times 10^{-4}</td>
<td>7.77 \times 10^{-4}</td>
</tr>
</tbody>
</table>

Table 6.3 Maximum IPCE data for Fig. 6.6c–f measured at two different biases
A series of pairs of internal standards consisting of undoped Fe and the Fe–Ti, Fe–Si, and Fe–Al binary systems with different drop densities (thickness) are also printed in the top row of the pattern as seen in Fig. 6.7.

IPCE values from the false color photocurrent data in Fig. 6.8a–d are summarized in Table 6.5. Again these values can be used to evaluate the relative responses when comparing the false color photocurrent scans since the brightest yellow shade was used to represent the maximum photocurrent response in each image. The highest overall IPCE data was obtained when the FeCl$_3$ precursor was printed three times with a drop spacing of 80 μm followed by drop spacings of 120, 40, and 160 μm (Figures 6.8b, c, a, and d, respectively). In Fig. 6.8a (FeCl$_3$ DS = 40 μm), the highest IPCE was observed at the Ti-rich area in this pattern (Fe$_1$Ti$_{0.0067}$Si$_{0.0067}$Al$_x$ top two rows) but the overall thickness of this library was too large for efficient carrier harvesting revealed by the dark squares in the centers of the squares. Fig. 6.8b (DS = 80 μm) has the brightest spots when the relative stoichiometry is near Fe$_1$Ti$_{0.0067}$Si$_{0.0067}$Al$_x$ that occurs in the center section of the library. In Fig. 6.8c (DS = 120 μm) the relative stoichiometry of Fe$_1$Ti$_{0.0017}$Si$_x$Al$_x$ in the second set of double rows appears to have the highest photocurrent. It is quite interesting to observe there is tendency in Fig. 6.8d for the highest photocurrents appearing in the Al-rich areas with stoichiometry of Fe$_1$Ti$_x$Si$_x$Al$_{0.095}$. Another interesting aspect is that in all of the libraries the pure α-Fe$_2$O$_3$ (top right two squares) and the binary internal standards (top row pairs 1–3) almost always have a lower photocurrent response than the ternary additions in their respective libraries.

---

**Fig. 6.7** Template used for the printed pattern in the Fe-Ti-Si-Al system where known amounts of the precursors are printed into individual 3 × 3 mm$^2$. The density of precursor drops and resulting stoichiometries for the various produced libraries are listed in Table 6.4.
Fig. 6.8 (a–d) False color photocurrent maps with 532 nm laser excitation of the Fe–Ti–Si–Al system with no applied bias in a 0.1-M NaOH solution (Top Left FeCl₃ precursor drop spacings: \( a = 40 \, \mu m, \ b = 80 \, \mu m, \ c = 120 \, \mu m, \ d = 160 \, \mu m \)). Note that the false colors are relative within a scan and not absolute values therefore comparisons of color between scans are not quantitative. Refer to Table 6.5 for quantitative comparisons and the top row of standard squares with only one dopant or no dopant (see text).

<table>
<thead>
<tr>
<th>Label</th>
<th>1D DS (( \mu m ))</th>
<th>Drops per ( \text{mm}^2 )</th>
<th>Relative stoichiometries (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>FeCl₃, 3 ( \times ) DS = 40 ( \mu m )</td>
</tr>
<tr>
<td>D</td>
<td>40</td>
<td>642</td>
<td>0.67</td>
</tr>
<tr>
<td>C</td>
<td>79</td>
<td>169</td>
<td>0.17</td>
</tr>
<tr>
<td>B</td>
<td>120</td>
<td>75</td>
<td>0.078</td>
</tr>
<tr>
<td>A</td>
<td>160</td>
<td>45</td>
<td>0.047</td>
</tr>
</tbody>
</table>

The iron precursor (FeCl₃) was over-printed three times

\( DS \) drop spacing
This indicates that there are synergistic effects of adding multiple dopants to the binary oxide semiconductor $\alpha$-Fe$_2$O$_3$. As stated in the introduction this empirical study does not provide information about which of the many possible mechanisms for photocurrent improvement provided by the impurity is actually occurring in the $\alpha$-Fe$_2$O$_3$ film. However, it does provide improved stoichiometries that should prove fruitful with more detailed studies of the actual electrical doping, morphological changes, carrier diffusion lengths and grain boundary and orientation effects attributable to the various impurities.

Due to the simplicity and low cost of the inkjet printing and screening protocols for discovering metal oxides capable of the photoelectrolysis of water, we embarked on a project to enlist multiple universities and even high schools in the search. We have produced the first beta printing and screening kits based on inexpensive components such as Lego Midstorms® kits, laser pointers and off-the-shelf inkjet printers that were distributed in the summer of 2008 to several universities with undergraduate summer research programs. The Solar Hydrogen Activity research Kit project (SHArK Project - www.thesharkproject.org) has recently expanded to over 20 sites world-wide and will be expanding to perhaps hundreds of sites in the near future. The SHArK Web site acts as a database of material combinations that have been investigated and as a communication tool for the various sites to exchange information. This “distributed research” project, initially funded by the Camille and Henry Dreyfus Foundation and now by the National Science Foundation, will lay the groundwork for a new generation of young scientists dedicated to helping solve the world energy problem.

### 6.4.3 Electrochemical Synthesis and Screening

To our knowledge McFarland et al. were the first to publish work on the combinatorial production of materials for water splitting. They demonstrated the electrochemical deposition of metal oxide compositions using robotics to electroplate and screen for water photooxidation activity individually created oxide materials using multiplexed photoelectrochemical cells [84]. Their work focused on variations in the electrochemical synthesis conditions, surface redox catalysts, control over structure [85] (by diverse surfactant and templating agents) and/or variable doping [86–88]. As illustrated in Fig. 6.9, the electrochemical deposition system utilizes a three-electrode electrochemical cell with coiled Pt and Ag wires serving as the
counter and reference electrodes, respectively, that can be moved with a computercontrolled x–y–z stage. Metal ligand systems (citrates, peroxo complexes, etc.) or electrolyte conditions (pH and relative concentrations) are carefully designed to produce similar deposition potentials for the species in solution as the precursors are electrodeposited onto the conductive substrate (FTO or Ti) surface and then connected as the working electrode for evaluation in the sample well. The desired voltage, with a necessary overpotential to ensure deposition of both species, is applied at each well until a specified amount of charge has passed before moving the scanning electrodes to the next well.

In this high-throughput configuration, each cell is manually filled using a robotic pipetter before electrodeposition. Subsequent calcination produces mixtures of known stoichiometries that were verified by XPS measurements coupled to Ar⁺ ion etching to examine the bulk and surface composition. Structural information was obtained with XRD. The screening system utilized for the electrochemically deposited samples incorporates measurements of flat-band potential, bias dependent photocurrent, and action spectra with a scanning photoelectrochemical probe [86]. Briefly, the probe steps from sample-to-sample across the library and an automated pump fills the probe with solution from a reservoir of stock electrolyte. An illuminated optical fiber targets each individual sample well and a three-electrode configuration is used to obtain photoelectrochemical data. A monochromator is placed between a chopped Xe lamp

Fig. 6.9 An automated electrochemical deposition system used for the production of metal oxides. The custom-made well plate shown has been designed for a 120-sample array, each sealed by “O-rings” underneath a patterned Teflon block, and each containing a unique electrolyte composition. Reprinted with permission from [86], copyright 2005 American Chemical Society
and the optical waveguide for the IPCE measurements. Using the multiplexed photoelectrochemical screening system, the measured photoactivity of large bandgap materials for oxygen production, including ZnO [86] and WO$_3$ [84], and hydrogen production, including Al$_2$O$_3$ [72], has shown improvement after doping with certain elements around particular stoichiometries and complete photoelectrochemical information is obtained for each sample.

### 6.4.4 Sol Gel Routes

The recent work of Maier et al. [88, 89] has also added to the arsenal of combinatorial production and screening strategies to the search for novel hydrogen-producing photocatalysts, particularly for producing doped mixtures of compounds with sol gel precursors such as titanium. The reported technique incorporates computer-controlled pipetting robots to quantitatively dispense stock solutions of precursors into an array of small vials. When the pipetting process is finished, the samples are placed in an orbital shaker for mixing and a consistent aging, drying, and calcination protocol is followed for each sample to produce catalyst powders. The powders are then manually ground and suspended in an aqueous methanol solution in small gas chromatography (GC) vials. After flushing the headspace within the vials with an inert carrier gas, the GC sample holder is illuminated with a visible light source and the direct production of hydrogen gas can be monitored.

Using this method, after several generations of optimization, a newly discovered material, Al$_{40}$Bi$_{40}$Pb$_{20}$O$_x$, was identified for future studies. There are several problems associated with this technique, including the use of powders and sacrificial agents such as methanol, both of which were previously discussed. Nevertheless, it would be interesting to deploy the materials discovered by this technique as thin film photoelectrodes to test their photoelectrolysis activity.

### 6.4.5 Other Thin Film Deposition Techniques

Although they have not yet been applied to the production and screening of photoelectrolysis materials, we will briefly mention some other approaches to produce metal oxide libraries using well-established thin film deposition techniques (meant here to include physical vapor deposition, sputtering, pulsed laser deposition, and molecular beam epitaxy). These techniques have been used for the production of combinatorial metal oxide libraries in the search for more effective luminescent materials [90, 91], transparent conducting oxides [92, 93], and dielectrics [94, 95]. It would presumably be straightforward to apply the same techniques to the production of material libraries to be screened for photoelectrolysis activity.
6.5 From Materials Search to Bulk Photoelectrodes

Identification of stable semiconducting oxide materials with appropriate band gaps and band positions is only the first step toward a viable photoelectrolysis system. In general, two factors conspire to limit the photoconversion efficiency of oxide semiconductors. First, the optical transitions in these materials are usually forbidden d–d transitions and so they have much lower absorption coefficients when compared to direct bandgap semiconductors. The result is that light penetrates deeply into the material before being absorbed to create an electron hole pair. Other indirect photovoltaic materials such as Si have a similar absorption problem but it is overcome by having very long carrier lifetimes, resulting in long carrier diffusion lengths, so that photogenerated electrons and holes are eventually separated. Carriers in metal oxide semiconductor d-bands usually have very low mobilities due to the limited overlap between the metal d orbitals, resulting in shorter charge carrier diffusion lengths and increased bulk recombination of photogenerated carriers. Therefore, even if we could use costly single crystals of a metal oxide semiconductor, the carrier collection would be restricted to the narrow space charge layer where only a small fraction of the solar photons will be absorbed. Luckily, these same two problems have been solved by the architecture of the nanocrystalline dye-sensitized solar cell, that has reached laboratory efficiencies of over 10% compared to sensitized single crystals of the same oxide semiconductor that show a quantum yield of at most 1–2% due to the low light absorption of the monolayer of dye. Simply making the dye layer thicker does not solve the problem as excitons produced away from the dye/semiconductor interface recombine before they are separated due to the short exciton diffusion lengths in the organic dye layer. The high surface area morphology works to increase the effective optical absorption for these systems as the multitude of dye monolayers absorb all the light and inject their electrons into the TiO₂ conduction band and the holes (or oxidized dye molecules) are quickly scavenged by a regenerator (reducing agent) in solution. A similar nanostructured morphology is probably necessary for photoelectrolysis with metal oxides in order to compensate for the low absorption and carrier mobilities characteristic of metal oxide semiconductors with forbidden d–d optical transitions.

Another illustration of the importance of morphology in a metal oxide semiconductor uses γ-Fe₂O₃ (hematite) as a photoelectrolysis material. Although iron (III) oxide is primarily used as a photoanode, it has been extensively studied to optimize its performance because it is inexpensive, easy to fabricate, has a bandgap (2.2 eV) that can absorb a good portion of the solar spectrum, and it is stable in aqueous solutions under illumination. High surface area nanostructured morphologies [96–98] of this material have been shown to more effectively harvest visible light relative to smooth films. For water cleavage, the best reported efficiencies (a solar-to-chemical conversion efficiency of 2.1%, based on the heat of hydrogen combustion), are achieved when used in a tandem configuration with a dye-sensitized solar cell.
The thickness of perpendicularly aligned \( \alpha\text{-Fe}_2\text{O}_3 \) sheets is between 15 and 25 nm allowing for carrier diffusion to the surface. This morphology appears to be a critical parameter, since samples produced by different methods with similar thicknesses, crystal structure, and identical optical characteristics do not perform nearly as well. The authors argue that this particular grain size, orientation and morphology (leaf-shaped sheets) are required due to the short hole diffusion length of only a few nanometers for \( \alpha\text{-Fe}_2\text{O}_3 \). This particular morphology affords facile transfer of holes to the electrolyte, before recombination occurs, while still shuttling electrons to the back contact. This morphology was produced using ultrasonic spray pyrolysis of their iron precursor solutions [86] to prepare their thin films. Addition of silicon produced further improvements in photoresponse again attributed to morphology changes rather than to electronic effects [84].

A nanostructured morphology also reduces the microscopic current density for water oxidation and reduction reactions resulting in lower overpotential losses than at a planar electrode. This will lower the electrocatalytic activity requirements for the semiconductor. In addition, many of the cell designs discussed herein or envisioned allow for thin nanostructured films deposited on a glass substrate with a transparent conducting layer. Such a configuration allows for back illumination, much like in the nanocrystalline solar cell. Back illumination of a photoelectrolysis electrode will reduce or eliminate the efficiency losses due to light scattering by the gas bubbles being evolved from the electrode surface.

### 6.6 Conclusions

We have surveyed the state of water photoelectrolysis using semiconductor electrodes and can conclude that there are currently no known suitable materials to efficiently and economically carry out this process. Due to the stability requirements we surmise that semiconducting oxide photoelectrodes are required. There are many potential oxide semiconductors that need to be prepared and tested for photoelectrolysis activity and therefore high-throughput combinatorial techniques are needed to discover promising materials. The first steps have been taken in the development of the combinatorial methods to solve this problem and these methods have been briefly reviewed. Optimization of the morphology of the semiconducting oxide thin film, probably with some kind of nanostructure, is required to overcome the limitations of light absorption and carrier transport in most oxide semiconductors. Mobilization of both experts and nonexpert researchers to discover and optimize new semiconductor materials is needed to produce the hydrogen needed to help move toward a clean energy economy based on renewable resources.
References


Chapter 7
Multijunction Approaches
to Photoelectrochemical Water Splitting

Eric L. Miller, Alex DeAngelis, and Stewart Mallory

7.1 Introduction

Hydrogen is an extremely valuable chemical commodity, not only in today’s industrial marketplace, but even more so in the emerging Green Economies so vital to the future of our planet and its people. Green futures will need to rely less-and-less on fossil fuels, and more-and-more on solar, wind, biomass, geothermal, and other renewable energy resources. Hydrogen is envisioned as one of the primary media for the storage and distribution of energy derived from this renewable portfolio. To achieve the vision, much work is needed in the development of new more practical technologies and infrastructures for hydrogen production, storage, delivery, and utilization. Nonpolluting technologies for large-scale hydrogen-production utilizing renewable energy are of particular importance.

Among the viable renewable hydrogen-production approaches, solar options, particularly solar-powered water-splitting, offer the best hope for large-scale renewable production with low carbon emissions. Several solar-to-hydrogen (STH) pathways are possible, including solar electrolysis, where electricity generated by photovoltaic (PV) or concentrated solar thermal (CST) power plants is used to drive electrolyzer systems, for example, commercial “alkaline” or “PEM” electrolyzers. Among a number of alternative, more direct conversion pathways, photoelectrochemical (PEC) water-splitting using semiconductor photoelectrodes or photocatalysts remains one of the most promising, though still elusive based on challenges in developing stable and efficient PEC materials, devices, and systems. Broadly speaking, a PEC system combines the harnessing of solar energy and the electrolysis of water into a single conversion system. As illustrated in Fig. 7.1,
sunlight shining on a suitable PEC semiconductor device immersed in a water-based solution can simultaneously drive water oxidation and hydrogen reduction electrochemical reactions to split water. This process evolves hydrogen and oxygen gases, effectively converting the solar energy into stored chemical energy.

The PEC approach is simple and elegant in concept, but is challenging to implement in practice since it relies on complex interactions involving sunlight, semiconductors, and liquid solutions. In a practical system, the semiconductor material must efficiently absorb sunlight and generate sufficient photovoltage to split water, while the semiconductor interface must be favorable to sustaining the hydrogen- and oxygen-gas evolution reactions. In addition, the PEC system needs to remain stable in solution, and must be inexpensive for compatibility with large-scale deployment. To date, no known semiconductor system achieves all of the above criteria — though some have come close. PEC devices based on multijunction III–V semiconductor technology have been demonstrated at the National Renewable Energy Laboratory (NREL) with impressive STH conversion efficiencies (i.e., the ratio of the chemical energy available in the generated hydrogen to the incident solar energy) exceeding 12% [1–4]. Unfortunately, these III–V devices have lacked long-term stability, and moreover, they utilize a number of prohibitively expensive materials as seen in the multilayered device structure illustrated in Fig. 7.2. Lower cost PEC devices based on thin-film semiconductors, also in multijunction configurations, have demonstrated stable hydrogen conversion efficiencies ranging from 3 to 8% [5–8], but progress toward higher performance in such devices has been slow. Before practical PEC hydrogen production can become a reality, key scientific advancements in the development of new PEC materials and devices are still needed.

Toward this end, the most advanced scientific techniques in materials theory, synthesis, and characterization are being brought to the table on an international scale, and powerful synergies among researchers in the PEC, photovoltaics, and nanotechnology fields are emerging in collaborative pursuit of the PEC quest [9].
In fact, PEC researchers often stake claim to PEC as the Holy Grail of hydrogen production [10, 11]. This may seem somewhat overly dramatic, though PEC is unique in offering an efficient low-cost approach for harnessing solar energy to produce high-purity hydrogen from water, at low operating temperatures, with no carbon emissions. The key to successful deployment will be in the identification and development of innovative materials systems, likely involving multijunction semiconductor configurations similar to that seen in Fig. 7.2.

Multijunction approaches offer some of the best hope for achieving practical PEC hydrogen production in the near term, but complicated materials and interface issues need to be addressed by the scientific community. This chapter explores the challenges and benefits of large-scale solar water splitting for renewable hydrogen production, with specific focus on the multijunction PEC production pathways. The technical motivation and approach in the R&D of multijunction PEC devices are considered, and examples of progress in laboratory scale prototypes are presented.

7.2 Solar-to-Hydrogen Conversion

Hydrogen is a valuable chemical commodity. Today, it is a vital component in the industrial processing of important chemical commodities, such as fuels and fertilizers. In the near future, it could take on new significance as a viable alternative to gasoline or electricity as an energy carrier for stationary power and transportation applications. On a commercial scale, STH conversion, specifically in the form of solar water-splitting, offers the best hope for renewable production with low carbon emissions. This approach for large-scale industrial deployment is truly renewable and sustainable considering the enormity of the solar resource and the abundance of water as a low-cost feedstock for hydrogen production. The “re-cycling” of water as a by-product of hydrogen fuel cell and combustion utilization schemes also factors into the sustainability of solar water splitting.
7.2.1 Value of Hydrogen

Hydrogen has become a hot topic in recent years, both in political and scientific circles. A national spotlight was cast on hydrogen in Former President George W. Bush’s State of the Union Speech in 2004, which featured such memorable quotes as: “America can lead the world in developing clean, hydrogen powered automobiles” and “With a new national commitment, our scientists and engineers will overcome obstacles to taking these cars from laboratory to showroom” [12]. Similar optimism has been expressed across the board by politicians, economists, and scientists, as exemplified in Jeremy Rifkin’s canonical work “The Hydrogen Economy” [13]. Still, there is by no means universal acceptance. From the opposing perspective, energy expert Joseph J. Romm in his book The Hype About Hydrogen asserts: “Neither government policy nor business investment should be based on the belief that hydrogen cars will have meaningful commercial success in the near- or medium-term” [14].

Whatever view you take on the future importance of hydrogen in our society, the world’s most abundant element is in fact an extremely valuable chemical commodity today. In the contemporary industrial marketplace, hydrogen is a high volume chemical with US production exceeding 20 billion kg annually [15]. Important industry uses include the production of chemicals, processing of materials, semiconductor manufacturing, generator cooling, and fertilizer production. In fact, the demand for hydrogen has been expanding significantly in recent years, especially in the fuels processing sector. As petroleum crude becomes dirtier, and as biomass feedstocks enter the fuels mix, growing amounts of hydrogen are being required for “hydrocracking” in the refining processes [16].

Hydrogen’s low density, high thermal conductivity, and strong chemical reducing properties make it ideal for the broad ranging industrial applications discussed above. To satisfy the expanding demand, current hydrogen production continues to rely primarily on established fossil-fuel technologies. Worldwide, over 95% of hydrogen is produced from natural gas, oil, or coal, with production relying largely on steam-methane reforming (SMR) with downstream purification processes such as pressure swing absorption (PSA) [17]. Of course, fossil fuels are vulnerable to dwindling availability and rising cost, and result in carbon emissions and other forms of environmental contamination. The current fossil-based industries for hydrogen production and utilization still suffer the risks and disadvantages of our current Fossil Fuel Economy. New, cleaner, and long-term approaches to hydrogen production and utilization are essential, not only to sustain current levels of demand, but also to accommodate future expanded Hydrogen Economy markets [13, 18–20].

In visions of a future Hydrogen Economy, hydrogen is an ideal energy carrier for the storage and distribution of renewable energy resources such as solar, wind, geothermal, and hydroelectric. Using fuel-cell or combustion engine technologies, hydrogen can be converted simply and cleanly to power or heat with no carbon emissions, and with water as the primary by-product. Unfortunately, although hydrogen is nature’s most abundant element, it exists primarily in strongly bonded chemical compounds, and extracting it involves difficult and energy-intensive processes. With current technologies, production, storage, and utilization of this
ideal energy carrier are all challenging. An enormous amount of technology and infrastructure development would be needed to attain a pure Hydrogen Economy. More realistically, a Green Economy will emerge comprising a broad portfolio of alternative energy sources and carriers, including hydrogen. As our present reliance on petroleum-based fossil fuels become increasingly difficult to sustain, both economically and environmentally, this will become inevitable. As a result, new, more distributed approaches to national and world energy management will take hold. Different locations rich in their own renewable resources will manufacture energy currencies such as electricity or hydrogen for large-scale distribution to the broader energy marketplace. Electricity is a key energy carrier today, and will remain so long into the future. Hydrogen, however, will also emerge in an important complementary role, providing important benefits in large-scale energy storage and long-distance distribution.

Hydrogen is indisputably valuable today, and will become increasingly valuable as an energy carrier with the future development of new renewable energy production and distribution infrastructures. In the process, new and improved technologies will emerge for the economical and environmentally friendly production, storage, delivery, and utilization of hydrogen. Production technologies using solar energy to split water are enormously attractive, motivating accelerated PEC research and development. Current research in multijunction PEC devices will be particularly important to near-term deployment of large-scale solar hydrogen production, as discussed in the following sections.

7.2.2 The Solar Resource

The true value of the solar resource as a renewable energy commodity becomes painfully apparent considering the current situation of oil. According to the US Energy Information Administration (EIA), the current worldwide oil in reserves totals approximately 213 trillion liters [21]. The amount of oil consumed globally as of 2008 was determined to be 5 trillion liters [22]. The projected oil consumption growth rate between 2004 and 2030 is expected to be approximately 1.6% [23]. Figure 7.3 shows projected oil consumption in relation to known oil reserves, both at current consumption levels and for a projected growth of 1.6% per year. As can be seen, even if the current consumption does not increase, the current oil reserves will become significantly depleted over the next 50 years.

New oil reserves will be found, yet the fact remains that this is a shell game of diminishing returns. Oil is a finite commodity, and before it runs out entirely, it will become increasingly energy-intensive and costly to extract from the earth and process into fuels. Creating a renewable alternative to satisfy our world’s energy appetite is likely to be the challenge of our Age. As Dr. Carl Safina aptly noted: “Whoever owns the new sources of energy will own the future economy” [24].

Our best hope for the future is the sun, the ultimate renewable energy resource, continuously bombarding Earth with about 180,000 TW of radiant power. This is enough to power about 10 quadrillion compact fluorescent light bulbs; over
10 million bulbs for every person on the planet! [25, 26]. Approximately 36,000 TW of the incident solar power strikes earth’s land masses where terrestrial-based solar-energy conversion plants could be constructed. To put this in perspective, our society on average consumes 13–15 TW, with projected increases up to 30 TW over the next 50 years, consistent with trend shown in Fig. 7.3 [27]. These energy consumption numbers are enormous, but still represent a small fraction of the sun’s influx of radiant power. Of course, the planet relies on the sun for many things, including sustaining plant-life and driving its weather patterns, so not all of the incident solar energy is available for mankind’s energy hunger. Still there is significant solar energy to spare. Using 10% solar energy conversion systems, 30 TW could be produced using only 0.25% of the land.

Unfortunately, it is not that easy. The large-scale conversion of solar energy to mechanical work or electricity is currently quite challenging and costly. At peak times of daylight, the solar intensity available for terrestrial conversion scales to approximately 1,000 W/m². Large collection areas and significant landmass would therefore be needed for commercial scale power production. Such expansive commercial deployment requires an enormous capital investment. For example, commercial photovoltaic technologies today can convert sunlight to electricity at efficiencies between 10 and 20%, at $2–5 per installed Watt [28]. A single Gigawatt plant would cost billions of dollars and span over 2,500 acres. As an added disadvantage, these expensive behemoths would be inoperative at nighttime and ineffectual under any significant cloud coverage.

There are practical difficulties and challenges, but the sun is still the most promising of renewable resources, and the most underutilized in modern society. Currently, less than 0.1% of the world energy production is from solar energy

Fig. 7.3  Projected exponential growth of oil consumption plotted for 0.0 and 1.6% growth rates given current global oil reserves.
plants, though this number is starting to increase [29]. Encouragingly, improved
technologies for solar energy conversion, storage, and utilization are emerging to
make their impact on the world energy scene. New and improved solar-to-electric
and STH conversion technologies are all poised to be part of the new energy mix.
Developing the most practical and efficient technologies to convert sunlight to
useable energy carriers such as electricity or hydrogen is a critical priority. With its
potential benefits, PEC hydrogen production remains an attractive conversion route
for large-scale solar hydrogen production.

### 7.2.3 Solar Conversion Pathways

In converting sunlight, whether to electricity or to hydrogen, fundamental thermo-
dynamic principles govern the energy conversion process. The sun radiates energy
at a black body temperature of 5,780 K, while the Earth’s black body temperature
is 300 K. The Carnot limit between these source and sink temperatures is readily
calculated as 95%, representing the amount of radiant energy can be converted into
other more useable energy forms [30]. Though this is encouraging, actual conver-
sion of sunlight to practical end-uses is always further limited by unavoidable losses
associated with the available energy conversion routes. Every added conversion
step in the process adds losses and reduces overall conversion efficiency.

A primary mechanism for converting solar energy on earth is the photoexcitation
of electrons in terrestrial matter. With this electronic excitation as an initial step, the
two basic routes for further energy conversion of the photoexcited electrons are
solar–thermal and solar-potential. In the solar–thermal routes, the energized
electrons thermalize to their surroundings, converting the energy to heat, which
can be converted further, for example, using heat engines to produce work. Such
routes are further restricted by lower Carnot limits based on the intermediate source
temperatures driving the heat engines. In the alternative solar-potential routes,
the elevated electrochemical potential of the energized electrons can directly
drive further conversion processes, for example, producing electricity or chemical
products. Thermal energy is not being converted to heat, so additional Carnot limits
are not imposed.

Based on these photoexcitation models for energy exchange, solar water-
splitting for hydrogen production can follow several different conversion routes,
as shown in Fig. 7.4. The solar–thermal route shown as process (A) in the figure is a
two-step process, with a photon-to-thermal energy conversion step followed by a
thermal-to-chemical (TC) conversion step. The other two-step process shown as
process (B) represents PV-electrolysis, where a photon-to-electric conversion step
is followed by an electric-to-chemical conversion process. The three-step process
(C) represents a concentrated-solar–thermal–electricity/electrolysis route, involv-
ing photon-to-heat, heat-to-electricity, and electricity-to-chemical conversion
steps. The final pathway depicted as process (D), representing a single-step direct
conversion from photon-to-chemical energy, is the PEC water-splitting process.
Although a direct route in theory, practical implementation, for example, with multijunction semiconductor devices, will entail thermal losses which need to be minimized for high-efficiency conversion. Other STH pathways are also possible, including photobiological routes [31, 32] and the ultrahigh temperature thermolysis route [33], all with their own sets of challenges.

A number of different pathways can contribute to the renewable production of hydrogen for future “Green Economies,” but economics will determine which ones will play the most important roles. In general, the most direct conversion processes, such as PEC water-splitting, could have some inherent economic advantages. In fact, PEC hydrogen production as a low-temperature single-stage process could be one of the front-running alternatives, contingent on development of sufficiently efficient, stable, and inexpensive material systems. Under the general umbrella of PEC solar hydrogen production, there are several possible plant strategies and configurations that could be practically feasible. These are described below.

### 7.2.4 Photoelectrochemical Hydrogen Plants

There are numerous hydrogen-production approaches utilizing solar energy for the PEC water splitting. With these different approaches, there are several possible commercial scale plant configurations [34, 35]. Ultimately, the practicality of any plant will be measured by the cost of the produced hydrogen. This will be determined by the efficiency, cost, and lifetime of the active PEC materials used, in addition to associated plant capital costs and operating expenses. As discussed in the following section, initial techno-economic analyses have been performed to compare and contrast economic advantages and disadvantages of different PEC plant concepts [35], but more comprehensive follow-on analyses will be needed,
especially as PEC conversion processes evolve and become more refined. Broadly speaking, PEC plant configurations fall into three categories: (1) combined PV-electrolysis systems; (2) photoelectrode-based systems; and (3) photocatalyst-based slurry systems. Many of the different single- and multijunction PEC materials and devices discussed in the following sections can be integrated into one or more of these plant types. Hydrogen production scales ranging from small-scale 1,500 kg/day for distributed production up to large-scale 50,000 kg/day central-plant production are being envisioned for all three categories. Some of the likely reactor schemes are illustrated in the following figures.

Figure 7.5a, for example, illustrates the most direct implementation of the “Combined PV-electrolysis” configuration. Commercially available photovoltaic (PV) panels are coupled with separate commercial electrolyzer units, such as alkaline or PEM electrolyzers; and appropriate power-conditioning equipment is utilized to load-match the processes. This is the clear path to near-term renewable solar hydrogen, but it is by no means inexpensive. Based on recent cost studies from the NREL, hydrogen production cost would exceed $10/kg for PV electricity cost at
The need for power conditioning equipment contributes to this high cost. Additional factors include the electrolyzer capital costs as well as the losses from lateral collection and transmission of electricity.

A variation of the PV-electrolysis approach is the integrated PV-electrolyzer unit shown in Fig. 7.5b. In such systems, the solid-state PV panels and large area electrolysis components are designed to be better matched in terms of both operational voltage and current density. Power conditioning is eliminated and transmission losses would be reduced. Based on the direct coupling to the solar flux rage, the electrolysis process would run at low current densities, for example, below 50 mA/cm², compared with rates over 1,000 mA/cm² in commercial electrolyzers. This relaxes the need for precious metal catalysts in the electrolysis process, and therefore reduces capital cost. Added complexity and cost of the dual-function panels for large-scale deployment is a significant tradeoff.

Unlike the systems in the “PV-electrolysis” category, the photoactive components in the “photoelectrode based systems” configuration are immersed in solution, and typically include one or more PEC solid/liquid rectifying junctions.

Figure 7.6a shows a classic two-electrode configuration with a photoelectrode immersed in aqueous solution and electrically interconnected to a metallic counter electrode. Sunlight incident on the photoelectrode energizes the water-splitting
process, promoting hydrogen evolution at the “cathode” surface and oxygen evolution at the “anode” surface. The PEC process entails both electron exchange in the external wire and ion exchange in solution. To minimize electronic and ionic conductivity losses, the counter electrode is placed in close proximity to the photoelectrode surface. In some cases, the counter electrode is a small catalyst wire placed right in front of the photoelectrode. This, however, introduces some optical shading loss; and it also requires a front-surface gas-separation membrane, adding further loss.

Figure 7.6b illustrates a single-electrode variation of the photoelectrode system comprising anode and cathode surfaces at opposite faces of a single monolithically integrated device. In this configuration, electron transfer is readily achieved through the body of the electrode, but pathways for ionic transfer in the electrolyte need to be provided around the electrode body to minimize solution conductivity losses. Electron transmission loss is reduced, as is system part count. This offers potential cost reduction.

The photoelectrode schemes in Fig. 7.6a, b require photoactive electrode materials and devices capable of spontaneously splitting water with solar illumination. As discussed at length in the following sections, this will most likely require the use of multijunction structures to generate sufficient voltage to sustain the process. Alternative external-voltage and chemical biasing strategies are possible in photoelectrode-based systems as illustrated in Fig. 7.6c, d. External biasing is specifically useful for single-junction photoelectrodes incapable of efficient unassisted solar water splitting. The added costs associated with the supplemental electricity needed in Fig. 7.6c or chemical feedstocks needed to maintain the pH gradient in Fig. 7.6d generally makes these approaches unattractive from a practical perspective.

In the photoelectrode-based systems, active semiconductor layers or thin-films are coated on fixed electrode structures which are immersed in solution. In contrast, the “photocatalyst”-based reactors utilize photoactive semiconductor particles free-floating as slurry in a solution bed. Figure 7.7a represents a conceptual single-bed photocatalyst reactor, comprising concentrations of one or more functionalized photocatalyst particles suspended in solution. In the example shown in Fig. 7.7a, the semiconductor particle would need to be capable of internally generating enough photovoltage to split water. In addition, specific surface sites on the particle would need to be functionalized, for example, with catalyst nanoparticles, to promote hydrogen and oxygen evolution and facilitate charge separation. An interesting variation of the single-bed photocatalyst system is the “Z-scheme,” which is somewhat analogous to the multijunction approach in the photoelectrode case. As described in more detail in later sections, the Z-scheme utilizes different photocatalysts for hydrogen and oxygen evolution, but the two particle types are mixed together in the same reactor bed with an ion mediator in solution to couple the reactions. Single-bed photocatalyst reactors are extremely simple, with perhaps the best potential for low cost at large scales. The obvious drawback is that potentially explosive mixtures of hydrogen and oxygen gas are generated in the same solution, necessitating safe and reliable gas separation mechanisms.
Separation technologies, such as pressure swing absorption, are commercially available, but would add complexity and cost.

Figure 7.7b illustrates a two-bed photocatalyst reactor. Similar to the Z-scheme, two separate photocatalyst particles are used, one for hydrogen evolution and the other for oxygen evolution. In the two-bed case, however, the different particle types are suspended in separate reactors which are connected by an ionic bridge-way to allow for mediator coupling. In the two-bed reactors, the hydrogen and oxygen gases are generated separately and safely, eliminating the cost of separations equipment. The drawback is the reduced hydrogen production efficiency inherent in this configuration. The ion bridge introduces conductivity loss; but more importantly, a larger
solar collection area is necessary for water splitting when hydrogen and oxygen reactions are stacked side-by-side, reducing conversion efficiency and increasing overall cost.

### 7.2.5 Techno-economic Analyses

The US Department of Energy (DOE) recently commissioned Directed Technologies Inc. (DTI) of Arlington, Virginia, to conduct a techno-economic evaluation of conceptual PEC hydrogen production systems [35] incorporating performance and processing cost feedback from the broader materials R&D efforts at DOE. The objective was to provide a basis for evaluating the long-term feasibility of large-scale PEC production technologies in comparison with other renewable approaches. The DTI study, which was completed in 2009, comparatively evaluated two photocatalyst particle bed configurations (compatible with the reactor configurations illustrated in Fig. 7.7) as well as two photoelectrode-based systems (compatible with the reactor configurations illustrated in Fig. 7.6). The four conceptual systems in the study were designated “Type-1” through “Type-4” as follows:

- **Type-1: Single-Bed Photocatalyst System**: A single electrolyte-filled reactor bed containing a colloidal suspension of PEC photocatalyst particles which produce a mixture of H₂ and O₂ product gases.
- **Type-2: Dual-Bed Photocatalyst System**: Two electrolyte-filled reactor beds containing colloidal suspensions of PEC photocatalyst particles, with one bed carrying out the oxygen evolution half-reaction, the other bed carrying out the hydrogen evolution half-reaction, and including a mechanism for circulating the ions between beds.
- **Type-3: Photoelectrode System**: A fixed array of planar PEC photoelectrodes immersed in an electrolyte reservoir, tilted toward the sun at local latitude angle, producing hydrogen at atmospheric pressure.
- **Type-4: Photoelectrode System with Moderate Concentration**: A PEC solar concentrator system, using reflectors to focus the solar flux at a 10:1 intensity ratio onto planar PEC photoelectrodes immersed in an electrolyte reservoir and pressurized to 300 psi.

In the DTI analyses, each of the modeled system types included a PEC reactor generating H₂ and O₂: photocatalyst-based on Types 1 and 2, and photoelectrode-based on Types 3 and 4. Each also included a gas processing system for compressing and purifying the output gas stream, in addition to ancillary balance of plant equipment. For each system type, the total cost of hydrogen production was projected for a 10-metric ton per day plant assuming technology-readiness of the PEC reactor. The baseline for readiness included targeted values of 10% for STH conversion efficiency and 5 years for PEC cell lifetimes. Sensitivity analyses around the baseline assumptions were performed to assess the effects on production...
costs of variations in the PEC performance and cost parameters. A summary of the
sensitivity analyses for the four system types is shown in Fig. 7.8.

As concluded in the DTI report, the hydrogen production costs are lowest for the
Type 1 and Type 2 photocatalyst systems, ranging from $1.50 to $5.25/kg H2.
These are encouraging results, consistent with the US DOE cost targets of $2–4/kg
H2 [38] for future central hydrogen production pathways. The report recognized,
though, that significant research and development work is still needed to demon-
strate particle-based PEC reactors operating at the baseline performance levels

Fig. 7.8 Hydrogen cost
sensitivity analyses for the
four reactor types evaluated
in the DTI techno-economics
report on Photoelectro-
chemical Hydrogen
Production. Projected
hydrogen costs as a function
of efficiency, lifetime, and
materials cost are shown.
assumed in the study. In contrast, the Type 3 photoelectrode system is currently the most mature of the concepts, with numerous prototypes demonstrated to date on the laboratory scale. The report concluded, however, that hydrogen production costs were the highest using this route, ranging from $6.14 to 18.75/kg H₂. The high costs in this case were dominated by substantial material and capital costs, similar to those plaguing the PV industry. The Type 4 photoelectrode system is a modification of Type 3 system with the addition of modest solar concentration (up to 10×) and moderate internal hydrogen compression (to 300 psi). Interestingly, these added features resulted in substantial cost reductions, specifically due to the reduction in PEC-electrode area, reduced amount of PEC materials needed, and fewer postprocessing compression stages. The resulting hydrogen production costs were reduced to $2.85–5.55/kg H₂, falling within the DOE targeted range.

Although preliminary, this study has shown that hydrogen production by PEC systems can be economically viable in both photocatalyst and photoelectrode configurations. However, key research challenge remains for attaining PEC material systems compatible with the baseline levels of 10% STH conversion efficiency and 5-year lifetime. As discussed in the following sections, multijunction approaches are likely to be critical for near-term realization of such performance levels. As research and development of PEC materials and devices progresses, more techno-economic analyses will be needed to further evaluate the tradeoffs between different large-scale PEC reactor approaches. Two of the most promising avenues appear to be photoelectrode systems using low-cost thin-film materials [39] and large-scale slurry bed reactors utilizing low-cost functionalized semiconductor particles [40]. The path forward for practical PEC solar hydrogen production needs to navigate all the economic as well as the scientific obstacles that remain.

7.3 Photoelectrochemical Water Splitting

In general, standard chemical processes involve interactions between chemical and ionic species. Electrochemical processes also involve interfacial interactions between ionic conductors, such as electrolytes, and solid-state electronic conductors, such as semiconductors. Involving even more complexity, PEC processes are exposed to light, so that optical photons can also interact with the electrochemical reactions. In semiconductor photoelectrochemistry, energetic photons of light typically create electron–hole pairs within the semiconductor that can participate in redox reactions at the semiconductor/electrolyte interfaces. The complicated set of fundamental electrochemical and solid-state optoelectronic principles that govern the behavior of such systems are well-documented in the literature [41–46]. Within the realm of photoelectrochemistry, the PEC water-splitting process itself presents a unique set of energetic and kinetic challenges [47–50]. Some useful simplifications to the problem, presented in the following sections, can provide a broad overview of PEC water-splitting, and help to highlight the loss mechanisms and conversion efficiency limitations inherent in PEC hydrogen production.
7.3.1 PEC Water-Splitting Reactions

Any light-driven water splitting process can be expressed as a generalized chemical reaction representing the addition of photonic energy to liquid water to split the H$_2$O molecule, evolving gaseous hydrogen and oxygen:

$$\text{H}_2\text{O} + 2\gamma \rightarrow \text{H}_2(\text{g}) + 1/2\text{O}_2(\text{g}) \quad \text{Net water-splitting reaction (7.1a)}$$

$$\Delta G^0 = + 237.18 \text{ kJ/mol} \quad \text{Standard Gibbs free energy. (7.1b)}$$

Here, $\gamma$ is the photon energy and $\Delta G^0$ is the standard Gibbs free energy. As indicated in (7.1), energetic photons, such as those in sunlight, can supply the energy to drive the uphill reaction converting the feedstock H$_2$O into hydrogen and oxygen gases. Specifically, the standard Gibbs free energy change of +237.18 kJ/mol is a quantification of the thermodynamic minimum energy needed for splitting water into the constituent gases at standard conditions of 25°C and 1 bar.

PEC implementations of light-driven water splitting utilize one or more photoactive semiconductor electrodes to convert the photon energy. In a PEC system, absorbed photons induce electron–hole pairs in the semiconductor bulk which can be separated and extracted to promote two “half reactions” in solution. Specifically, photogenerated electrons drive the hydrogen evolution half-reaction (HER) at the cathode interface, while photogenerated holes drive the oxygen evolution half-reaction (OER) at the anode interface. Both half reactions must be sustained simultaneously, coupled by their exchange of electrons in the solid state, and ions in solution. In real systems, multiple reaction pathways and reaction steps are possible at both anode and cathode, and physical conditions shape if and how the process will proceed. As one example, the half-reactions for water splitting can be written explicitly in terms of the H$^+$ ionic exchange between anode and cathode:

$$2\gamma \rightarrow 2e^- + 2h^+ \quad \text{Photon-induced electron-hole pair generation (7.2a)}$$

$$\text{H}_2\text{O} + 2h^+ \rightarrow 2\text{H}^+ + 1/2\text{O}_2(\text{g}) \quad \text{OER: anodic water-oxidation half-reaction (7.2b)}$$

$$2\text{H}^+ + 2e^- \rightarrow 2\text{H}_2(\text{g}) \quad \text{HER: cathodic H$^+$ reduction half-reaction (7.2c)}$$

$$\text{H}_2\text{O} + 2\gamma \rightarrow \text{H}_2(\text{g}) + 1/2\text{O}_2(\text{g}) \quad \text{Net water-splitting reaction (7.2d)}$$

$$\Delta G^0 = + 237.18 \text{ kJ/mol} \quad \text{Standard Gibbs free energy (7.2e)}$$

$$\Delta V^0_{\text{rev}} = \Delta G^0 / nF = 1.23 \text{ V} \quad \text{Standard reversible potential (7.2f)}$$

$$V_{\text{op}} = \Delta V^0_{\text{rev}} + \eta_a + \eta_c + \eta_\Omega + \eta_{\text{sys}} \quad \text{Operating voltage with overpotential loss. (7.2g)}$$
Here, $e^-$ is an electron, $h^+$ is a hole, $F$ is Faraday’s constant, and $n (=2)$ is the number of electrons exchanged. $V_{\text{op}}$ is the operational voltage, $\eta_a$, $\eta_c$, $\eta_\Omega$, and $\eta_{\text{sys}}$ are overpotentials associated with anode, cathode, ionic-conductivity, and system losses, respectively.

Alternatively, the half-reaction set can also be written in terms of $\text{OH}^-$ ionic exchange:

\[
\begin{align*}
2\gamma & \rightarrow 2e^- + 2h^+ \quad \text{Photon-induced electron-hole pair generation} \quad (7.3a) \\
2\text{OH}^- + 2h^+ & \rightarrow \text{H}_2\text{O} + 1/2\text{O}_2(g) \quad \text{OER: half-reaction} \quad (7.3b) \\
2\text{H}_2\text{O} + 2e^- & \rightarrow 2\text{OH}^- + \text{H}_2(g) \quad \text{HER: half-reaction} \quad (7.3c) \\
\text{H}_2\text{O} + 2\gamma & \rightarrow \text{H}_2(g) + 1/2\text{O}_2(g) \quad \text{Net PEC water-splitting}. \quad (7.3d)
\end{align*}
\]

The corresponding expressions for the Gibbs free energy, the standard reversible potential and the operating voltage are equivalent to those given by (7.2e–g).

The reaction pathways depend on the nature of the electrodes, the electrolyte, and the light-induced energy levels. For example, the half-reaction formulations expressed in terms of $\text{H}^+$ tend to dominate in acidic electrolyte solutions, while the $\text{OH}^-$ formulations can dominate in basic solutions, specifically due to the predominance of one ionic species over the other under the different conditions [51]. In either case, it is important to emphasize that each of the half-reactions described in (7.2) and (7.3) is simplifications of more complex multistep electrochemical reaction pathways, and that competing or parasitic reactions are also possible [52, 53]. Without parasitic reactions and losses, the Gibbs free energy of photo-driven water splitting is +237.18 kJ/mol, and the standard reversible potential of PEC water splitting is 1.23 V, as indicated in (7.2) and (7.3). This is independent of the dominant pathways involved in achieving the net reaction. All practical systems will involve losses, such as those represented in the various overpotential terms in the equations; and these losses can be significantly different for different reaction pathways. Including overpotential losses, the energy/voltage needed to split water will inevitably exceed 1.23 V.

### 7.3.2 PEC Water Splitting Losses

Losses incurred during PEC water splitting include the overpotential losses at the anode and cathode interfaces, the ionic conductivity overpotential loss associated with ion transport in the electrolyte, as well as other solid state and balance of system losses. The anode and cathode overpotential losses include the effects of activation energy, kinetics, and mass-transport of the multistage half reactions. These can be substantial, commonly several tenths of volts, with more severe loss
for oxygen evolution, which is the more complex and less facile reaction [54]. In addition to the electrode overpotential losses, the electrochemical losses due to ionic conductivity in the solution can also be severe, depending on electrode geometry and spacing, in addition to solution properties. The splitting of pure water, for example, is particularly difficult, since the ionic conductivity, typically less than 0.05 S/m range is prohibitively low. Weak acid or alkaline solutions with conductivities exceeding 10 S/m are typically used to compensate, although this creates a more corrosive environment for the electrodes.

On top of the losses at the interfaces and in solution, there are also significant losses in the solid state of any PEC system. While the ions in solution are migrating through the aqueous media and reacting at the interfaces, the solid-state electrons–holes need to be exchanged between the anode and cathode through a conductive pathway (such as an interconnecting wire in two electrode schemes and the conducting substrate in monolithic single electrode configurations), and therefore some ohmic loss is inevitable. Additionally, there can be severe losses in the bulk semiconductor and solid-state interfaces of the electrodes. These include optical absorption losses, electron–hole pair recombination losses, and interface recombination losses, among others, which also degrade system performance.

The PEC water splitting process is a delicate balancing act, where photon-energized electron–hole pairs under the right conditions can simultaneously drive the electrochemical half-reactions. In steady state, the anodic and cathodic half-reactions need to be sustained at the same reaction rate, or a charge build-up will occur, impeding and eventually stopping the entire process. A similar situation exists with the charge carriers in the solid state. The anodic half-reaction generates two electrons (i.e., “consumes” two holes), while the cathode half-reaction consumes two electrons. These electrons must be shuttled from anode to the cathode via electrical current and steady state cannot be maintained if anode and cathode reaction rates are not the same.

As a result of all the electrochemical, solid-state losses and any balance of system losses occurring at steady state, water splitting cannot occur at the bulk reversible potential of 1.23 V. The operating voltage for water-splitting must exceed this value to compensate for all the losses. In practice, water electrolysis systems typically require operating voltages of 1.6–1.9 V, depending on gas production rates [55, 56]. In other words, to drive the water-splitting process including all solution and electrode losses, the absorbed photons must induce sufficient electrochemical potential to the electron–hole pairs to generate photopotentials in excess of approximately 1.6 eV. The photoelectrolysis balancing act can be set into motion ONLY if this photopotential requirement is met. As discussed in the following sections, it is the photopotential requirement that may necessitate the use of multijunction approaches for efficient PEC water splitting. If and only if the conditions are right to sustain the water splitting process, the hydrogen evolution will then be proportional to electron consumption rate at the cathode, which represents a photocurrent. This is critical to the understanding of STH conversion efficiency.
### 7.3.3 PEC Solar-to-Hydrogen Conversion Efficiency

During steady-state operations, the solid-state shuttling of charges between anode and cathode represents a photon-induced current, or photocurrent. This photocurrent is integrally tied to the hydrogen-producing performance of the PEC system. Specifically, the PEC half reactions from the previous section illustrated that two electrons are consumed in the evolution of one \( \text{H}_2 \) molecule. This results in a hydrogen production rate that is half the rate of electron flow, in other words, half the photocurrent:

\[
R_{\text{H}_2} = \frac{I_{\text{ph}}}{2e} = \frac{(J_{\text{ph}} \times A)}{2e},
\]

where \( R_{\text{H}_2} \) is the hydrogen production rate \( (s^{-1}) \), \( I_{\text{ph}} \) is the photocurrent \( (A) \), \( e \) is the charge of an electron \( (C) \), \( A \) is the area of the illuminated photoelectrode \( (m^2) \), and \( J_{\text{ph}} \) is photocurrent density \( (A/m^2) \).

The chemical STH conversion efficiency of a solar-based hydrogen production system is defined as the ratio of the useable chemical energy in the generated hydrogen gas to the total solar energy delivered to the system [57–61]. For steady-state operations, this is equivalent to the ratio of the power output to the power input of the system. This power ratio can be expressed as:

\[
\frac{P_{\text{out}}}{P_{\text{in}}} = \frac{(\text{hydrogen production rate}) \times (\text{hydrogen energy density})}{\text{solar flux integrated over illuminated area}}.
\]

Using the hydrogen production rate from (7.4) and the Gibbs energy as the useful energy density of the hydrogen, the STH efficiency for a PEC system can be written as:

\[
\text{STH(\%)} = \frac{\Delta G}{N_A} \times \frac{R_{\text{H}_2}}{P_{\text{solar}} \times A} = \frac{\Delta G}{N_A} \left( \frac{J_{\text{ph}} \times A}{2e} \right) = \frac{\Delta G}{2N_A P_{\text{solar}}} \times J_{\text{ph}},
\]

where \( N_A \) is Avogadro’s number and \( P_{\text{solar}} \) is the solar flux of energy in W/m\(^2\). The first ratio in (7.6) is generic for any STH production system, while the second term is derived specifically for PEC hydrogen processes using the production rate from (7.4).

The solar flux of radiant energy is comprised of a broad spectrum of energetic photons, which are quantized particles of light, each with discrete energy content. Figure 7.9a shows the standard atmosphere-filtered global solar spectrum (AM1.5G) indicating the range of photon energies and the distribution of energy transmitted by these photons [62]. This AM1.5G spectrum is also frequently presented as a function of photon wavelength, as shown in Fig. 7.9b.
Integrating the AM1.5G solar flux yields a total power density of $\sim 1,000 \, \text{W/m}^2$ for sunlight. Using this value for $P_{\text{solar}}$ in (7.6), the STH conversion efficiency can be related directly to the photocurrent density of a PEC system:

$$\text{STH}_{\text{AM1.5}} \, (\%) \approx \begin{cases} 0.123 \times J_{\text{ph}} & \text{for } J \text{ in (A/m}^2) \\ 1.23 \times J_{\text{ph}} & \text{for } J \text{ in (mA/m}^2) \end{cases} \, (7.7)$$

Fig. 7.9 AM1.5 global solar irradiance spectrum: (a) as a function of photon energy and (b) as a function of photon wavelength, indicating the ultra-violet, visible, and infra-red regions of sunlight.
Equation (7.7) explicitly states the linear relationship between conversion efficiency and photocurrent density calculated under AM1.5G solar illumination. The use of the Gibbs free energy in these equations reflects chemical energy in the hydrogen that can be retrieved using an ideal fuel cell. This in effect, calculates the lower heating value (LHV), which is standard in practical comparisons between different fuels [63].

It is clear from (7.7) that the PEC photocurrent density is the key determining factor for STH efficiency. This is in contrast to solid-state photovoltaic cells, which need to be operated at the maximum power point (i.e., maximum product of photocurrent and photovoltage) for the best solar-to-electric conversion efficiency [64]. It is a subtle but important distinction that PEC cells should be operated at maximum photocurrent for best hydrogen-production performance. It is the bandgap-limited saturated photocurrent density of a semiconductor that ultimately constrains the hydrogen production rate. For peak STH efficiency, sufficient useable photopotential must be generated in the device to drive the photocurrent as close as possible to the saturation limit. With all the built-in losses in the solid state, in solution, at the interfaces, and in the balance of system, this can be a challenge. In fact, due to limits of useable photopotential generation in a single junction, it may be necessary to resort to multijunction schemes to drive the PEC system into the maximum photocurrent region. There are inherent trade-offs and limitations as well as benefits to this approach. These are explored in detail in the following sections.

7.4 Single Junction PEC Semiconductor Devices

In a semiconductor-based PEC system, photons are absorbed in the semiconductor and the photogenerated charge carriers are separated, and then extracted by the interfacial HER and OER reactions. In photoelectrode-based systems, the charge separation is largely due to the rectifying nature of the semiconductor/electrolyte junction. The charge separation details are somewhat different in the photocatalyst systems [65], but the basic absorption and extraction mechanisms at the solid/liquid interface are the same. In either case, hydrogen production performance is strongly influenced by both semiconductor material properties and interfacial characteristics. An examination of junction formation and operation in the case of photoelectrode configurations reveals much about the key performance parameters and conversion efficiency limitations.

7.4.1 Photoelectrode Junction Formation

In PEC photoelectrode systems, the semiconductor/electrolyte interface can form a rectifying junction, similar to the solid-state pn junctions or Schottky diode junctions used in solar cells. Such rectifying junctions exhibit built-in electric fields
capable of separating photogenerated charge carriers (i.e., electron–hole pairs) created by absorption of photons in the semiconductor bulk. In PV cells, this charge separation mechanism drives photocurrents to produce electricity, while in the PEC case, the charge separation can drive the HER and OER half-reactions for water splitting. In both cases, illumination creates extra photoexcited electron–hole pairs which need to be separated and extracted before they recombine in order to be effective in the energy conversion process. Extraction of the photogenerated charge carriers with elevated electrochemical potentials in effect converts the solar energy to electricity or hydrogen in PV or PEC systems, respectively. Excellent sources of information are available detailing semiconductor material properties, solid-state junctions and solar cells, and treating the fundamentals of rectifying junction formation and behavior [26, 50, 66–70].

Formation of a rectifying PEC junction at a semiconductor/electrolyte interface follows similar principles, but key concepts from both solid-state physics and electrochemistry need to be combined for a complete description of the process. For reference, there is a wealth of literature on fundamental electrochemical principles [41–46, 50], which are useful, especially in conjunction with the previous citations covering semiconductor physics. Of particular interest to PEC studies are the models developed by Gerischer, which establish important connections between the in-solution electrochemical potentials of electrons and solid-state Fermi levels [47–49]. Using the Gerischer models, descriptions of semiconductor/electrolyte junctions follow closely the solid-state junction analogies. Photoanodes using n-type semiconductors form PEC junctions similar to a Schottky barrier or an np+ junction, and internal electric fields are set up to drive photogenerated holes toward the electrolyte interface promoting oxygen evolution. In contrast, p-type photocathode junction formation is analogous to a solid-state pn+ device, and the internal fields promote the injection of photogenerated electrons at the interface to drive the HER [50]. The field directions for both n-type and p-type junctions are shown explicitly in the following illustrations of junction formation.

Specific steps in semiconductor/electrolyte junction formation are illustrated in Fig. 7.10 for both the photoanode (7.10a) and photocathode (7.10b) cases. In Fig. 7.10a (i), the n-type semiconductor and electrolyte are shown in thermal equilibrium prior to contact for the photoanode case. At equilibrium, the Fermi level in the semiconductor \( (F_n) \) is close to the conduction band (CB), and the Fermi level in solution \( (F_s) \) falls between the redox (reduction/oxidation) levels for hydrogen reduction \( (H^+/H_2) \) and water oxidation \( (H_2O/O_2) \). After contact, the electrode and electrolyte Fermi levels align to reach thermal equilibrium, as shown in Fig. 7.10a (ii). Since the initial electrode Fermi level is higher than the electrolyte Fermi level, free electrons in the n-type semiconductor will migrate to the solid–liquid interface exposing positively charge fixed donor sites. The electrons form a surface charge layer at the interface which induces a thin Helmholtz double layer in the electrolyte. A depletion-region, also known as a space-charge region (SCR), forms where free charge carrier diffusion is counter-balanced by the built-in electric field generated by the fixed charges flanking the junction. The charge distributions including the fixed space charges in the solid-state electrode and the Helmholtz layer charges in solution are shown in Fig. 7.10a (iii). Typically, Helmholtz layers are on the
order of a few nanometers, compared with several micrometers for the semiconductor space-charge region.

For the p-type photocathode case, the equilibrium charge distribution process is the same, but as shown in Fig. 7.10b the charges and band bending are reversed. As illustrated in Fig. 7.10b (i) and (ii), the initial migration of holes toward the interface during contact equilibration creates a space charge region of fixed negative acceptor charges, reversing the direction of the built in electric field compared with the photoanode cases. In both cases, the thermal equilibrium is disturbed upon exposure to light, and illuminated operations need to be described in terms of “quasi-equilibrium” statistics, which generally remain valid for AM1.5G insolation as well as moderate levels of concentrated sunlight [66, 67, 70].

### 7.4.2 Illuminated Response

Under illumination, the electron–hole pairs generated by photon absorption in the vicinity of the space-charge regions can be separated by the built-in electric
fields. Some of the photocharges will be successfully extracted in the water splitting process, while the remainder will recombine, losing energy as radiation or heat. In the photoanode case, the photogenerated holes will be driven by the built-in field toward the electrolyte interface, where, with the appropriate energetics and kinetics, they can drive the OER oxidation reaction. At the same time, photoexcited electrons are driven toward the electrode’s back contact, where they can be extracted to solution at counter-electrode surfaces appropriately conditioned to promote the HER. For photocathodes, the charge separation supplies photogenerated electrons to the hydrogen evolution process at the electrolyte interface, and drives photogenerated holes to an OER counter electrode surface. Under these operating conditions, the initial thermal equilibrium has been disturbed by the energy influx of the light, and the band descriptions with Fermi levels are no longer valid. Under conditions of quasi-equilibrium, a useful alternative description can be developed in terms of “quasi-Fermi” levels.

Figure 7.11 illustrates PEC junction response to solar illumination for both photoanodes (Fig. 7.11a) and photocathodes (Fig. 7.11b) using the quasi-Fermi level descriptions. For both configurations in Fig. 7.11, the photoelectrode is immersed in solution, and the back contact is connected by external wiring to a counter electrode also in solution. In addition, the \( \eta_a \) and \( \eta_c \) overpotentials for the OER and the HER in solution have been added to the reversible potential of the redox system to stress the minimum voltage requirement to sustain water splitting.

With the absorption of sunlight in the semiconductor bulk, the original Fermi levels split into separate quasi-Fermi levels for electrons (\( F_e \)) and holes (\( F_h \)) [47, 66, 67, 70] resulting from the excess concentration of photogenerated electron–hole pairs over the equilibrium populations. For the photoanode case in Fig. 7.11a, the excess hole population significantly alters the minority carrier distribution with respect to the equilibrium populations, while the excess electrons barely affect the majority carrier numbers. As a result, the hole quasi-Fermi level shifts substantially in contrast to insignificant change in the electron quasi-Fermi level. Conversely, in the photocathode case in Fig. 7.11b, there is a significant shift in the electron quasi-Fermi level, but little hole quasi-Fermi level shift. In both cases, near the back contact away from the effects of illumination, the two quasi-Fermi levels converge back to a bulk equilibrium Fermi level.

The back contact potential is tied to the counter electrode potential, which under operating conditions is coupled to the back half reaction. As an important result, the quasi-Fermi separation in the semiconductor determines the useable energy for driving the front half reaction. For the illuminated photoanode device shown in Fig. 7.11a, the electron potential at the counter electrode is sufficiently high to drive the hydrogen-reduction half-reaction, including the \( \eta_c \) loss. Simultaneously, the quasi-Fermi hole energy at the solution interface is sufficiently low to drive water-oxidation, including the \( \eta_h \) loss. This configuration, in consequence, is capable of sustaining the net PEC water-splitting process, driven by useable energy in the quasi-Fermi split.
7.4.3 Useable Photopotential

The concept of the quasi-Fermi level split is extremely important to how much useable photopotential a semiconductor device can generate, specifically in relation...
to the bandgap energy. In crystalline silicon photovoltaic cells, for example, the bandgap energy is 1.1 eV, while single-junction open-circuit voltages typically range between 0.6 and 0.7 V [26]. Thermodynamically, semiconductor band diagram representations reflect the internal energy of electrons and holes, not the useable energy. Electricity can be extracted from a PV cell only at potentials below the open-circuit voltage, which is typically 50–70% of the semiconductor bandgap energy. The output voltage limit can be increased using higher bandgap cells, but as a tradeoff, fewer solar-spectrum photons would be absorbed, limiting the saturated photocurrent.

The concept of useable potential is even more important in the case of PEC water-splitting devices. The useable potential of a PEC device must be sufficiently high to drive both half reactions in addition to the overpotential losses. This can typically require over 1.6–1.9 V. To achieve this in a single junction device, semiconductor bandgaps over 3.0 eV are typically necessary, which severely restricts optical absorption, saturated photocurrent, and therefore, STH conversion efficiency.

The bandgap position relative to the redox potentials of the electrolyte must also be considered. It is often reported that semiconductors with bandgaps “straddling” the redox potentials will be able to photosplit water in single junction configurations. While this is a “necessary” condition, it is by no means a “sufficient” condition. The band positions of semiconductor interfaces in aqueous solutions have been reported [5]. Figure 7.12, for example, shows the band edge positions with respect to the redox levels for hydrogen reduction and water oxidation for several commonly studied PEC semiconductor materials. Of these materials, titanium dioxide, zinc oxide, and gallium nitride all have bandgaps greater than 3 eV, and band edges straddling the redox levels. None of these, however, have demonstrated spontaneous PEC water splitting. This further drives home the point that the real energetics and useable photopotential under illumination are dictated by the quasi-Fermi levels, and not by the band edges. In addition, it is important to stress that all the overpotential losses directly subtract form useable potential. As a result, there are many situations where a bandgap straddling the redox potentials will fail to split water.
To illustrate this further, two additional hypothetical devices are illustrated in Fig. 7.13a, b representing photoanode and photocathode junctions, respectively, that are incapable of solar water splitting. In both cases, the conduction and valence band-edges clearly straddle the redox levels, even including the overpotential.
losses. In Fig. 7.13a, however, the hole quasi-Fermi level at the PEC interface is too high, while in Fig. 7.13b, the electron potential in the counter electrode is too low and thus in either case PEC water-splitting is not sustainable.

Interestingly, these “dysfunctional” photoelectrodes could, in theory, be combined in tandem to achieve sufficient useable potential for water splitting. This is illustrated in the combined band-diagram representation seen in Fig. 7.14, where the two electrodes have been interconnected to each other, instead of to counter electrodes. This utilization of multiple junctions to enhance useable photopotential is a powerful tool for PEC water-splitting. As discussed in the following sections, multijunction stacking can offer important benefits, but can also introduce new system losses that limit overall STH conversion efficiency. In all single- and multijunction configurations, loss minimization in the device and system designs becomes critical to hydrogen production performance and efficiency.

7.4.4 Loss Minimization

There are numerous losses in PEC devices and systems that can limit STH conversion efficiency, and which need to be minimized. These include losses in solution,
at interfaces, and in the semiconductor bulk. The optoelectronic losses in the semiconductor photoelectrode, for example, need to be as low as possible. In single-junction absorbers, photons with energies below the semiconductor bandgap cannot be absorbed or converted. Photons with energies exceeding the bandgap are absorbed, but at rates dependent on the allowed transitions in the semiconductor. Direct bandgap materials, for example, absorb more efficiently than indirect bandgap materials. In all semiconductors, photogenerated electron–hole pairs rapidly thermalize to band-edge energy levels within picoseconds, losing energy to heat. Large bandgap semiconductors generate little photocurrent due to poor absorption, while small bandgap semiconductors can suffer from low conversion efficiencies due to high thermalization losses. At the band-edge energy states, the electron–hole pairs can often survive for several microseconds before recombining, and they must be separated and transported to electrochemical interfaces for extraction during this time for effective energy conversion. This separation is assisted by the electric fields set up by charge distributions in the semiconductor and at the solid/liquid interface. Defects in the bulk and at the interface can adversely affect the separation fields, and also result in poor mobility for charge transport. If wide absorption widths are needed, for example, in indirect semiconductors (i.e., the absorber needs to be relatively “thick”), the charge transport losses can be severe. Thinner devices with back-surface optical reflection may reduce the loss.

To minimize overpotential loss due to interfacial charge extraction and electrochemical product formation, the interface conditions need to be optimized. Ideally, charge is extracted via the water-splitting half-reaction at the solid/liquid interface. The extraction process can be slowed or completely inhibited by poor energetic alignment or poor surface kinetics at the photoelectrode or counter electrode surfaces. Parasitic or corrosion reactions competing with the water-splitting reactions can also result in substantial loss; though surface treatments can be employed to kinetically and/or energetically favor water-splitting over the parasitic processes. Surface incorporation of nanoparticle catalysts is one approach, though light blockage due to such particles must be avoided. Since PEC water-splitting is a low-current density process, typically operating below 20 mA/cm², nonprecious metal catalysts can be used. Additionally, nanostructuring of electrode surfaces can increase effective surface area for enhanced charge extraction, although this can also lead to higher surface recombination loss.

The electrolyte is an important factor determining stability, efficiency of the charge-extracting reactions, and the electrochemical byproducts. Splitting seawater, for example, is a challenge since it is difficult to electrochemically suppress the production of chlorine gas from the Cl⁻ ions [71]. During PEC water-splitting, the evolved hydrogen or oxygen gas must be efficiently removed from the photoelectrode surface to avoid mass-transport losses in the surface reactions, and to minimize adverse optical effects. Surfactants added to the electrolyte have been successful in promoting rapid bubble formation and dissipation. In solution, ionic conductivity losses tend to be a larger problem. High electrolyte concentrations can be used to minimize this loss, but the tradeoff is in higher corrosivity.
Photoelectrode geometry and counter-electrode proximity are critical parameters to the redistribution of ions. In some geometries, gas separating membranes are needed, introducing further ionic transport loss.

Overall, the key condition for PEC hydrogen production is that the quasi-Fermi levels under solar illumination generate sufficient useable photopotential to drive the redox reactions for water splitting after all the solid-state, interfacial and solution overpotential losses have been taken into account. This challenge has severely limited efficiency in all single-junction PEC devices to date.

### 7.4.5 Single Junction Efficiency Limits

It is difficult to achieve high STH conversion efficiencies in single-junction PEC photoelectrode systems with inherent electrochemical and solid-state losses. Large bandgap semiconductor materials would be needed to generate sufficient useable photopotential. This, however, can severely limit photon absorption, therefore reducing photocurrent and STH conversion efficiency. The bandgap tradeoff between photopotential and photocurrent is particularly detrimental for single junctions. For example, if the minimum water-splitting potential, including the redox separation and overpotentials, amounts to 1.6 V, then the required quasi-Fermi level separation is also 1.6 eV. Even for high-quality semiconductor materials, the quasi-Fermi level separation can only achieve 50–70% of the bandgap level [72]. In this case, the minimum bandgap for the onset of photoelectrolysis would be 2.7–3.2 eV. For increasing levels of hydrogen production, the system losses will also increase, including the current-dependent surface overpotentials as well as electrical and ionic conductivity losses. Even higher bandgaps would then be required.

An upper bound for STH conversion efficiency can be established as a function of semiconductor bandgap based on fundamental optical absorption limits [50, 73]. Figure 7.15 plots the maximum attainable AM1.5G photocurrent densities in a semiconductor as a function of bandgap based on optimal light absorption and carrier extraction. The derivation assumes that every photon in the solar spectrum with an energy exceeding the bandgap will create an electron–hole pair, and that all of these electron–hole pairs are converted to photocurrent. For bandgaps greater than 3.2 eV, the photocurrent density is limited to approximately 1 mA/cm². This places an upper limit of 1.23% on the STH efficiency, as calculated using (7.4). The STH values corresponding to the maximum achievable photocurrents are listed in parentheses on the right vertical axis of Fig. 7.15, but these only apply to junctions capable of spontaneous water-splitting. Strictly speaking, any photoelectrode system incapable of sustaining photoelectrolysis will not produce photocurrent and thus have 0% STH efficiency.

Fundamental thermodynamic limits for the conversion efficiency of single- and multijunction semiconductor PEC devices have also been established in numerous studies [74–81]. The thermodynamic single-junction limit under ideal conditions, i.e., including thermalization losses but with no overpotential losses, has been
calculated to be approximately 30% STH. With recovery of some thermalization losses through advanced multi-exciton device schemes, this limit is modestly increased to 32% STH [74]. More thorough analyses have been reported which include overpotential losses in addition to the optical and thermalization losses. These have indicated that the best single junction PEC cells for AM1.5G water-splitting will likely be limited to STH efficiencies below 12%, even using the highest quality semiconductor and catalyst materials currently available [82].

7.4.6 Single Junction Examples

To date, the only demonstrations of single-junction water-splitting have utilized very high bandgap materials such as SrTiO$_3$ and KTaO$_4$ [83, 84]. Based on poor photon absorption, the demonstrated STH values have been well below 1%, consistent with the values in Fig. 7.15. Commonly studied semiconductor PEC materials have included iron oxide (Fe$_2$O$_3$), tungsten trioxide (WO$_3$), and titanium dioxide (TiO$_2$). The theoretically achievable photocurrent densities highlighted in Fig. 7.15 look encouraging, especially those for iron oxide. However, none of these materials can develop enough useable photopotential under sunlight to split water, not even the TiO$_2$ with a bandgap over 3.0 eV. To enhance the photopotential, these and other promising semiconductors can be incorporated in multijunction PEC schemes. This is consistent with the multijunction device enhancements that are well known in the PV community. For PEC devices, the same approach can be taken to enhance the photopotential and also to increase the absorption efficiency, as described in the following sections.
7.5 Multijunctions

Multiple-junction semiconductor devices for solar energy conversion are well-known, especially to the PV research community. The use of multijunction approaches for high photovoltaic conversion efficiency is the basis of current technologies for space applications, and one of the major cornerstones for future Third Generation PV devices [85]. In theory and practice, integration of multiple junctions in stacked solar cells enhances the absorption of photons in the solar spectrum [64, 85, 86]. In fact, multijunction III–V solar cells with efficiencies greater than 40% are commercially available [87, 88]. Efficiencies in PEC solar conversion devices can also be enhanced by improved spectral absorption in multijunction configurations. More importantly, multijunction PEC devices can develop increased photopotentials, better meeting the voltage demands of water splitting. The tradeoffs are reduced photocurrent, which adversely affects hydrogen production rates, and added device complexity. Overall, a PEC multijunction device design must strike the right balance to maximize STH conversion efficiency without adding too much cost. The multijunction approach, with its benefits, limitations, and tradeoffs are further discussed in the following sections.

7.5.1 Multijunction Stacks

The concept of stacking multiple photorectifying cells is illustrated in Fig. 7.16. Though the figure specifically depicts a triple cell stack, the illustration can be generally extended to represent any number of stacked cells. Each cell in the stack represents an independent photoconverting junction, such as a semiconductor pn junction, or a semiconductor/electrolyte PEC junction. Electrically and optically the cells are stacked together in a series-connected fashion. Incident sunlight on the “top” cell is partially absorbed, generating a photovoltage and photocurrent in the process. The remaining filtered light reaches the “middle” cell, where further absorption and photoconversion takes place. Any remaining sunlight can be absorbed and converted in the “bottom” cell. Since the cells are series-connected, the individual photovoltages will add, but the net current will be the minimum of the individual cell photocurrents. This is because the net current through the device must be continuous, and will be bottlenecked by the minimum component cell current. Any excess carrier generation in the other cells will be lost internally to recombination. Since the individual cells are all powered by some fraction of the incident solar photons, particularly the buried cells, they operate at reduced photocurrents compared with full-sun levels. As a result, overall device current is reduced by the bottleneck effect. Typically, then, a multijunction device based on a given semiconductor material system will generate lower photocurrent compared with a single-cell device of the same material. In general, multijunction devices offer higher voltage with lower current. As illustrated in later sections, over-all conversion efficiencies can be enhanced due to the extended range of photons effectively absorbed using multiple bandgaps in an integrated device.
7.5.2 Optical Considerations

Semiconductor photorectifying cells absorb photons in the solar spectrum, and can convert some of the absorbed photon energy into electricity or chemical products, such as hydrogen. Different semiconductors are specifically sensitized to different parts of the solar spectrum dependent on optoelectronic properties related to bandgap and absorptivity. In general, the macroscopic light absorption process in a material follows the Beer–Lambert Law [89], here expressed as a function of photon wavelength, $\lambda$:

$$I_\lambda(x) = I_{0,\lambda} e^{-\alpha_\lambda x}.$$  

For each photon wavelength $\lambda$ in (7.8), $I_{0,\lambda}$ is the intensity of the incident flux (e.g., in W/m$^2$), $\alpha$ is the absorption coefficient (in m$^{-1}$), and $x$ is the length of the optical pathway through the material (in m). $I_\lambda$ is the unabsorbed flux intensity emerging out the back of the material, as illustrated in Fig. 7.17.
Important information about the spectral-sensitivity of a given semiconductor material is incorporated in the absorption coefficient $a$, which is strongly dependent on material bandgap and allowable energy transitions. Figure 7.18, for example, depicts typical absorption coefficient curves for four different crystalline semiconductors: (1) gallium indium phosphide (GaInP$_2$), a direct band transition semiconductor with bandgap 1.86 eV (corresponding to a photon wavelength of 667 nm); (2) gallium arsenide (GaAs), a direct band transition semiconductor with

![Diagram of Beer-Lambert Law](image1)

**Fig. 7.17** Illustration of the Beer–Lambert Law for wavelength-dependent photon absorption in a material with absorption coefficient $a$

![Absorption Coefficient Spectra](image2)

**Fig. 7.18** Absorption coefficient spectra of GaInP$_2$, GaAs, Ge, and Si

Important information about the spectral-sensitivity of a given semiconductor material is incorporated in the absorption coefficient $a$, which is strongly dependent on material bandgap and allowable energy transitions. Figure 7.18, for example, depicts typical absorption coefficient curves for four different crystalline semiconductors: (1) gallium indium phosphide (GaInP$_2$), a direct band transition semiconductor with bandgap 1.86 eV (corresponding to a photon wavelength of 667 nm); (2) gallium arsenide (GaAs), a direct band transition semiconductor with
bandgap 1.43 eV (868 nm); (3) germanium (Ge), an indirect transition material with bandgap 0.66 eV (1,880 nm); and (4) silicon (Si), an indirect transition material with bandgap 1.12 eV (1,110 nm). For all four cases, the absorption coefficient vanishes at the gap energy, consistent with nonabsorption of photons with energies below the bandgap.

For photons with energies above the bandgap, the direct transition materials absorb more efficiently. This is reflected in the steeper slopes and correspondingly higher values for $\alpha$ in GaInP$_2$ and GaAs near their band energies, specifically compared with Ge and Si. Solar energy conversion devices based on direct-transition materials such as GaAs can be made as thin as a few micrometers and still absorb most of the available photons. In contrast, silicon solar cells need to be made much thicker, typically 100 $\mu$m or more to absorb most of the sunlight. A key point relevant to multijunction device design is that photon absorption and photocurrent generation in the individual component cells can be tailored by the selection of semiconductor bandgap and absorption properties. They can also be adjusted by varying cell thickness, as indicated in the Beer–Lambert law.

Another important parameter describing the optoelectronic response in a photoactive semiconductor device is the quantum efficiency (QE) (more precisely, the external quantum efficiency or EQE), which is critical to both single- and multijunction device designs. For each wavelength of light, the “external” QE expresses the percentage of photons incident on the device surface that results in an extracted electron–hole pair. This parameter is therefore an accurate measure of the device’s ability to generate photocurrent, and convert solar energy. The QE ratio and the corresponding relationship to photocurrent generation in a given semiconductor device can be expressed as:

$$\text{QE} = \frac{\#\text{charges/s}}{\#\text{photons/s}} = \frac{I/e}{P_{\text{total}}(\lambda)/E_{\text{photon}}(\lambda)},$$

(7.9)

where $e$ is the elementary charge, $I$ is the current, $P_{\text{total}}(\lambda)$ is the total power of incident light at a wavelength $\lambda$, and $E_{\text{photon}}(\lambda)$ is the energy of a single photon at wavelength $\lambda$.

The quantum efficiency is closely related to the absorption coefficient, but incorporates additional information regarding the efficiency of extracting photocharges created in the absorption process. Figure 7.19 shows representative QE spectra for GaInP$_2$, GaAs, and Ge devices, with sensitized regions of the AM1.5G included above. It is evident that the different semiconductors are sensitized to specific portions of the solar spectrum. Ge exhibits a broad bandwidth in its spectral sensitivity, but there is a relatively low level of solar photon flux across this bandwidth. In contrast, both GaInP$_2$ and GaAs have more narrow bandwidths, but each is efficient at converting photons in high-flux portions of sunlight. It is important to emphasize that QE is dependent not only on absorption and charge separation properties of device, but also on device operating conditions, such as voltage biasing, which alter
charge distributions and electric fields. These operating conditions need to be specified in reported QE curves to be useful in device design.

The solar energy conversion in a single bandgap semiconductor device is ultimately limited by the spectral width of the quantum efficiency curve. Also, since photogenerated electron–hole pairs rapidly thermalize to near band-edge energy states, energy conversion is most efficient for near-bandgap wavelengths. Absorbed photons with energies in excess of the bandgap will lose this excess to thermalization, reducing the conversion efficiency of light at the corresponding photon wavelength. Referring to the QE curves in Fig. 7.19, individual GaInP₂, GaAs, and Ge devices, for example, can convert photons in the 300–700, 700–900, 900–1,800 nm ranges, respectively, with optimal conversion near the 667, 868, and 1,800 nm bandgaps, respectively. By extension, a triple-junction stacked device using all three materials could absorb over a broader range of wavelengths from 300 to 1,800 nm, with optimal conversion efficiency at all the component cell bandgaps of 667, 868, and 1,800 nm. This illustrates an important avenue for enhanced efficiencies in multicell device schemes, relevant to both PV and PEC systems. The strategy has been studied extensively in the PV research and development community, and “full spectrum” absorber devices based on stacked multiple junctions are
envisioned as a future pathway to ultrahigh solar-to-electric conversion efficiencies [90]. As discussed in following sections, photovoltaic efficiencies in triple-junction III-V devices have already exceeded 40%.

### 7.5.3 Current and Voltage Relationships

The external quantum efficiency spectrum indicates the efficiency of photogenerated charge extraction and current generation as a function of photon wavelength. The QE specified at the device operating condition coupled with the incident photon flux provide sufficient information for determining the photocurrent of the device. Specifically, the theoretical carrier extraction rate is determined by integrating the product of QE and flux over the light spectrum. The photocurrent density is then equal to the extraction rate scaled by the ratio of electron charge/photon energy. The photocurrent equation can be explicitly written as an integral over photon wavelengths in the incident irradiation:

\[ J_{ph} = \int_{\text{irradiance}} \Phi_{\lambda} \left( \frac{e \lambda}{hc} \right) \frac{\text{QE}_{\lambda}}{C_{18}/C_{19}} d\lambda \]  

\( (7.10) \)

In (7.10), \( J_{ph} \) is the photocurrent density (e.g., in A/m\(^2\) or mA/cm\(^2\)) at device operating conditions, \( h \) is Planck’s constant, and \( c \) is the speed of light. Wavelength-dependent parameters include the incident irradiance flux, \( \Phi_{\lambda} \), and the quantum efficiency, \( \text{QE}_{\lambda} \), which is specified at the operating conditions.

As an example of \( J_{ph} \) determination from QE and solar flux, Fig. 7.20 illustrates the QE response at short-circuit conditions for a typical monocrystalline silicon PV single-junction cell along with the corresponding responsive regions of the AM1.5G spectrum. Integrating QE according to (7.10) using AM1.5G as the input irradiance yields 35 mA/cm\(^2\), consistent with the short-circuit photocurrent levels of commercial single-junction silicon PV devices [64]. As seen in Fig. 7.20, silicon, with a bandgap of 1.12 eV, has a broad spectral response for absorbing sunlight. Efficient energy conversion with low thermalization loss in this single-junction device, however, is limited to absorbed photons with energies near the bandgap energy (wavelengths near 1,100 nm). As a result, compared with the PV efficiencies exceeding 40% in multijunction III–V devices, efficiencies are limited to about 20% using single-junction silicon.

The useable photovoltage (equivalently, the photopotential) generated in a semiconductor device under illumination can be represented by the split in the quasi-Fermi levels as described in Sect. 7.4. This voltage is a function of semiconductor bandgap, but it is always less than the internal energy separation between conduction and valence bands represented by the bandgap. In PV cells, the maximum useable potential, expressed by the open-circuit voltage, is typically 50–70% of the bandgap energy [64, 91]. For example, silicon has a bandgap of 1.1 eV, and
typical open circuit voltages of single junction crystalline silicon cells are around 0.6 V. Similarly, high-quality GaAs has a bandgap of 1.4 eV, and device open circuit voltages of approximately 1.0 V. As the photoconversion device is asked to deliver current to an external load, the operating voltage will start to decrease below the maximum potential. The photovoltage and photocurrent levels are integrally related. Specifically, the maximum photovoltage is logarithmically related to the current-generating capability of the device, as derived from junction analysis [64, 92]. For photovoltaic cells, this is expressed as a relationship between the open-circuit voltage, $V_{oc}$, and the short-circuit current-density, $J_{sc}$:

$$V_{oc} = V_T \ln\left(\frac{J_{sc}}{J_0} + 1\right),$$

where $V_T$ is the temperature-dependent thermal voltage of the device, and $J_0$ is the device’s saturation current density.

The current and voltage relationships coupled with the optical parameters described in the previous section for single material systems provide powerful insights into the voltage, current, and conversion efficiency properties of single junctions, and also of stacked semiconductor device configurations. Another important performance parameter representing the quality of photorectifying junctions is the “fill factor,” or FF. The FF expresses how severely voltage will drop below the maximum useable potential under current loading. This is illustrated in Fig. 7.21, showing photoconversion device performance curves with “good” and “poor” fill...
factors. In PV cells, the maximum power point, found at the “knee” of the performance curve, is greatest for high fill factors. For PEC hydrogen production cells, it is necessary to maintain a high photopotential to drive the water splitting reactions while maintaining high photocurrent to maximize production rates. In other words, it is desirable to operate a PEC cell at maximum possible current with minimum voltage drop below the device’s maximum useable potential. A good fill factor is also critical in this situation. A high solar conversion efficiency in PV and PEC devices requires material systems and devices with maximum-photovoltage, maximum-photocurrent and fill factor all optimized.

The current and voltage relationships for individual single-junction cells can be combined to model of multijunction devices. The generic triple-junction scheme shown in Fig. 7.16, for example, can be modeled by the set of equations summarized in Table 7.1. These equations specifically highlight the dependencies of device photovoltage and photocurrent on the individual cell absorption coefficients and quantum efficiencies. This multijunction equation set is explicit to a three-cell device, but readily generalized to devices with any number of stacked cells. Additionally, the relationships expressed are applicable whether the individual cell junctions are photovoltaic, PEC, or a combination. If the device stack is designed to generate electricity, the PV conversion efficiency is determined at the “maximum power point,” as illustrated in the response curves of Fig. 7.21. Conversely, if the stack is designed to utilize solar energy for purposes of splitting water, the PEC conversion efficiency depends on the maximum achievable electrochemical photocurrent. In either case, optimizing voltage, and even more importantly, maximizing current is desirable for achieving high efficiency. Some key take-away points from the relationships in Table 7.1 relating to multijunction device design and optimization include the following:

- **Photocurrents are decreased in multijunctions**
  - Photocurrent in a component cell is dependent on the physical and optoelectronic properties of that cell, such as thickness, bandgap, and QE, but also on the available photon flux.
Photon flux available to a component cell is filtered, therefore reduced, by absorption in preceding cells in the stack. This optically couples photocurrent generation in all the cells.

The output current of the stack is limited to the smallest of the photocurrents produced by any of the constituent cells, and optical coupling effects between cells become important in device design.

Photovoltages are enhanced in multijunctions

Photovoltages of the components cells are additive in the stacked device, providing higher useable photopotential.

Photovoltage in a component cell is logarithmically dependent on the photocurrent, so current reduction due to filtering will decrease voltage, but not significantly.

Optics are important

Available flux in a component cell is determined by light absorption in previous cells, but light reflection is also important. Minimizing reflection at the front surface and at each interface while maximizing back-surface reflections to redirect light back through the device can significantly enhance efficiency.

Fill factor is important

Although fill factor is not represented explicitly in the equation set, optimal fill factor is important in each cell to ensure maximum output photocurrents with minimal loss in photovoltage.

### Table 7.1 System of equations relating voltage, current, and spectral responses in photoconversion cells and stacks

<table>
<thead>
<tr>
<th></th>
<th>Cell 1</th>
<th>Cell 2</th>
<th>Cell 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absorber bandgap ((E_g) (eV))</td>
<td>(E_{g1})</td>
<td>(E_{g2})</td>
<td>(E_{g3})</td>
</tr>
<tr>
<td>Thickness ((t) ((\mu m))</td>
<td>(t_1)</td>
<td>(t_2)</td>
<td>(t_3)</td>
</tr>
<tr>
<td>Absorption coefficient ((\alpha) ((um^{-1}))</td>
<td>(\alpha_1(\lambda))</td>
<td>(\alpha_2(\lambda))</td>
<td>(\alpha_3(\lambda))</td>
</tr>
<tr>
<td>Photon irradiance ((\Phi) (mW/cm(^2)/nm))</td>
<td>(\Phi_1 = 1) sun</td>
<td>(\Phi_2 = \Phi_1 e^{-\alpha_1 t_1})</td>
<td>(\Phi_3 = \Phi_2 e^{-\alpha_3 t_3})</td>
</tr>
<tr>
<td>Spectral response (CR) (C/J)</td>
<td>(SR_1 = \frac{q}{E_g(\lambda)} QE_1)</td>
<td>(SR_2 = \frac{q}{E_g(\lambda)} QE_2)</td>
<td>(SR_3 = \frac{q}{E_g(\lambda)} QE_3)</td>
</tr>
<tr>
<td>Short-circuit current ((J_{ph})) (mA/cm(^2))</td>
<td>(J_{ph1} = \int_{\lambda_{g1}}^{\infty} \Phi_1 SR_1 d\lambda)</td>
<td>(J_{ph2} = \int_{\lambda_{g2}}^{\lambda_{g1}} \Phi_2 SR_2 d\lambda)</td>
<td>(J_{ph3} = \int_{\lambda_{g3}}^{\lambda_{g2}} \Phi_3 SR_3 d\lambda)</td>
</tr>
<tr>
<td>Open-circuit voltage ((V_{oc})) (mV)</td>
<td>(V_{oc1} = V_T \ln\left(\frac{J_{ph1}}{J_0}\right) + 1)</td>
<td>(V_{oc2} = V_T \ln\left(\frac{J_{ph2}}{J_0}\right) + 1)</td>
<td>(V_{oc3} = V_T \ln\left(\frac{J_{ph3}}{J_0}\right) + 1)</td>
</tr>
<tr>
<td>Total device (V_{oc})</td>
<td>(V_{oc} = V_{oc1} + V_{oc2} + V_{oc3})</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total device (J_{ph}) Minimum ((J_{ph1}, J_{ph2}, J_{ph3}))</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Using these general points as guidelines, it is possible to develop practical designs for highly efficient multijunction devices. There are still a number of general design challenges for practical implementations that need to be considered.

### 7.5.4 General Design Challenges

In effective designs, maximum current in a multijunction photoconversion device is achieved by “current-matching” the individual cells in the stack [93]. For current-matching, the component cell materials and thicknesses can be tailored to ensure all component cell photocurrents are approximately the same. The goal is to minimize internal current losses. Since current must be continuous across the entire device, the lowest component cell photocurrent will bottleneck the output. If current excess is generated in any of the other component cells, it will be lost through internal recombination. If all junction currents are the same, this recombination loss is minimized, and net current is maximized. As an alternative viewpoint of current matching, the cells in the stack are essentially “splitting” all of the incident solar energy photons that can be used for photocurrent generation. Again, since the cell with the lowest current dictates the device current, an equal split will be optimal. Since photocurrent generation is sensitive to the incident flux of photons as well as physical and optoelectronic properties of all the cells, the process of designing an optimal current-matched device is complex. A general approach involves stacking semiconductors with decreasing bandgaps. Top layers absorb higher energy photons while transmitting lower energy photons that are subsequently absorbed by following layers of the device. After a component cell’s bandgap and quantum efficiency have been established, cell thickness can be adjusted to alter the number of absorbed photons, and tune the photocurrent.

An important challenge in the design of multijunction devices is the limited selection of availability of semiconductor materials with appropriate bandgaps that offer high efficiency through low defect densities. Examples of PV material systems that have been successful or have shown promise for multijunction configurations include the III–V crystalline materials, amorphous hydrogenated silicon films, and copper chalcopyrite thin-films [86, 94, 95]. In each case, bandgap tuning is available through modification of alloy composition. A broader class of low-cost thin-film PEC material systems for multijunction water-splitting applications is under investigation, including novel metal oxides and oxi-nitride compounds covering a broad range of bandgaps [96–99]. In the PEC junction case, there are additional material challenges associated with the potential and band-edge requirements for water splitting, as well as stability in solution.

In addition to the material selection challenge, interface design for low optical and defect recombination losses in stacked cells is complicated, and highly dependent on the materials being interfaced. In crystalline semiconductor multijunction devices, lattice-matching between adjacent layers is critical for avoiding defects and significant interfacial recombination losses [100, 101]. In polycrystalline and
amorphous thin film material systems, stress and strain at the various film interfaces must be minimized to avoid defects and even possible delamination of the device. In semiconductor material systems where heavy doping can be readily achieved, thin “tunnel junctions” are the preferred method of forming low-loss connection between adjacent cells [101]. Precise process control is essential in successful tunnel junction implementations. In other material systems where tunnel junction formation is not possible, the use of conductive transparent oxides can be used to interface adjacent cells, but optical and contact resistance losses can be introduced. To retain the efficiency advantages of multijunctions, all interface losses need to be minimized.

Broader challenges exist in the processing and manufacture of monolithic multijunction designs. In crystalline, polycrystalline, and amorphous semiconductor processing, precisely controlled process conditions are essential to the formation of device quality materials; and these conditions are highly material specific. Developing a process sequence to successfully stack numerous semiconductor layers can be extremely difficult. During the process, fabrication conditions at later stages, particularly high-temperature processes, can adversely affect the quality of all preceding layers. Low temperatures are advantageous for process compatibility and for lower manufacture cost, but it is often difficult to obtain the necessary material quality. The trade-offs boil down to enhanced performance vs. added complexity and manufacture cost. In PEC water-splitting applications, the increased photopotentials developed in multijunction devices is an added benefit to the mix. In the PV industry, the multijunction approach has been a clear winner in applications when the highest conversion efficiencies are absolutely required, despite higher cost.

### 7.5.5 Photovoltaic Multijunctions

The multijunction approach for achieving enhanced conversion efficiencies in PV devices have been studied for several decades. Since the early 1980s, solar-to-electric conversion efficiencies in production scale multijunction PV cells have been increased from just over 15% to over 40% today [86, 102]. The most efficient present-day multijunction photovoltaic cells are based on high-quality, and high cost, crystalline III–V material systems. For example, commercial cells have been on the market for several years, primarily for space applications, utilizing the GaInP$_2$, GaAs, and Ge materials discussed in Sect. 7.5.2. These specific materials were selected based on their collective ability to absorb photons over a broad range of the solar spectrum, as was seen in Fig. 7.19. GaInP$_2$, with a bandgap energy of 1.85 eV, absorbs short wavelength ultraviolet and visible photons in the solar spectrum; GaAs, with a bandgap of 1.42 eV, absorbs near-infrared light; and Ge, with a bandgap of 0.67 eV absorbs the lower energy photon energies in the infrared. In the device fabrication process, layers of crystalline GaInP, GaAs, and Ge layers are monolithically grown on a Ge substrate using epitaxial methods [103]. The completed triple-junction device scheme, containing over 20 layers, is illustrated in Fig. 7.22.
Several critical design features were necessary to enable high performance in these devices. For example, current matching among all three component cells was essential for maximizing photocurrent output. To avoid losses due to interface defects, the semiconductor layers in this configuration had to be lattice matched. Additionally, low-loss electrical contact between adjacent cells was achieved using optimized tunnel junctions, comprised of a ministack of highly doped layers producing an effective potential barrier for both minority carriers. Furthermore, a broadband dual-layer dielectric stack antireflection coating was employed. The resulting multijunction device achieves PV efficiencies over 34% under concentrated sunlight [104]. More recent variations on the device scheme utilizing a Ga0.44In0.56P/Ga0.92In0.08As/Ge stack have achieved over 40% PV efficiency under concentrated sunlight with open-circuit voltages over 3.0 V and operating voltages near 2.75 V [87]. These efficiencies represent world records in the field of photovoltaic solar energy conversion, but this comes at a price. The crystalline III–V material systems are prohibitively expensive, especially for any large-scale deployment.

As a lower cost alternative, multijunction devices incorporating thin-film semiconductors have also been developed. As an example, triple junction thin
film solar cells based on amorphous silicon/germanium alloys have been developed with photovoltaic efficiencies over 10% and with open-circuit voltages up to 2.3 V [54]. The fundamental design strategies for optimizing efficiency in these lower cost devices are similar to those in the III–V designs, including current matching, tunnel junction optimization, and minimization of optical losses. In a typical device, the thin-film silicon–germanium alloy bandgaps are adjusted to approximately 1.8, 1.55, and 1.4 eV in the top, middle, and bottom cells, respectively, to achieve current matching and broad spectral response. Also in the design, front-surface antireflection coatings are incorporated, while back surface reflector layers are used to reflect photons back into the device. Compared with single-junction devices using amorphous silicon–germanium material systems, the multijunction approach fully utilizes more of the solar spectrum. Higher conversion efficiencies are therefore achieved, but again at the expense of added device complexity and cost. Rigorous efforts in loss minimization are critical to ensure that the enhanced performance in the PV multijunctions offer worthwhile advantages over simpler single-junction devices. Efforts to achieve even higher efficiencies in economical thin-film multijunction technologies are on-going: not only in amorphous silicon devices, but also in alternative, enhanced absorbers based on bandgap-adjustable copper chalcopyrite alloys. Similar efforts are also underway to develop efficient multijunction water-splitting PEC schemes incorporating a broader class of low-cost semiconductor materials, such as metallic oxide alloys. Additional design considerations apply in the PEC case.

7.5.6 PEC Multijunction Design Considerations

Multijunction device designs can combine PV cells to improve solar-to-electricity conversion efficiency, or they can incorporate a combination of PV and PEC cells to enhance solar hydrogen production via water splitting. In both cases, the stacked junctions can be beneficial, providing photovoltage enhancement, as well as the broadened solar absorption width. The primary tradeoff is added device complexity and cost. In the PEC case, an additional tradeoff is reduced photocurrent, which limits hydrogen production rates. To make multijunctions worthwhile, the PEC device design must strike the right balance to maximize conversion efficiency and minimize expense. The general design considerations discussed in preceding sections still apply, including current matching and interface loss minimization. There are, moreover, additional considerations specifically applicable to a PEC component cell, which could be incorporated into a multijunction stack as photoanode (OER electrode) or a photocathode (HER electrode). These considerations include the following:

1. Low loss semiconductor PEC electrode materials with appropriate bandgap, absorption properties, and device quantum efficiency are needed for efficient photon absorption and charge extraction.
2. Sufficient useable potential must be developed (i.e., overall split in quasi-Fermi levels for holes and electrons) across the stack under illuminated operating
conditions to overcome overpotential losses and drive the OER and HER reactions.
3. Suitable band-edge position at the semiconductor/electrolyte interface need to be established to facilitate the proper alignment of the quasi-Fermi levels under illumination.
4. High surface activities for the OER/HER reactions need to be established to minimize the interface overpotential losses.
5. High surface passivation, based on energetics and/or kinetics, needs to be established to mitigate corrosion and all other parasitic reactions.
6. There needs to be enough flexibility in selection of semiconductor bandgap, absorption properties and thickness to allow for current-matching in the stack design.
7. Losses associated with electrolyte conductivity and oxygen/hydrogen product gas management (related to relative electrode placement, ionic pathways between electrodes, possible membrane separators, and other related design features) need to be minimized.

Design optimizations, for all the component cells and for the integrated stack, are essential for realizing enhanced efficiency in multijunction devices. Decades of research and development in PV multijunctions has resulted in solar-to-electric conversion efficiencies exceeding 40% in triple junctions, approaching theoretical limits. In contrast, PEC multijunction research is at a far less mature level. Though STH conversion efficiencies exceeding 10% have been demonstrated in best PEC tandem devices based on high-quality III–V crystalline semiconductor materials, these “First Generation” devices lack long-term stability and affordability. More stable and affordable “Second Generation” thin-film PEC multijunction devices have been demonstrated, but only in the 3–5% STH range of performance levels. The following section examines the theoretical limits for any “First-,” “Second-,” or even “Third-” Generation multijunction systems, with specific emphasis on attainable STH efficiency.

### 7.5.7 Multijunction Efficiency Limits

As with the single-junction PV or PEC devices, efficiency bounds can be placed on multijunction configurations based on optical absorption limits. Figure 7.23 is a two-dimensional extension of Fig. 7.15 for tandem devices, where maximum photocurrent, and the corresponding STH levels for PEC devices, are calculated as a function of both top- and bottom-junction bandgaps. The assumptions included in the derivation of this graph are as follows:

1. The top cell absorbs all photons with energies exceeding the top-cell bandgap.
2. The bottom cell is illuminated with the top-cell-filtered light, and absorbs all of the remaining photons with energies exceeding its bandgap.
3. For every photon absorbed, an electron–hole pair is generated and extracted as photocurrent.

4. The total photocurrent of the stack is then simply the smaller of the two photocurrents produced by each individual cell.

5. For the case of PEC cells, the STH efficiencies, indicated in brackets in the plot, are calculated using (7.7). These ONLY apply to systems with sufficient photopotentials to split water.

The results from Fig. 7.23 are definitely encouraging for the viability of tandem photovoltaic electricity production as well as PEC hydrogen production. Despite the reduced photocurrents inherent in the tandem arrangement, a wide range of bandgap combinations yield high enough photocurrents for STH conversion efficiencies over 10%. An important key is to identify the possible bandgap combinations capable of the necessary photopotential levels for water splitting. The useable potential of a single cell ranges from 50 to 70% of the semiconductor bandgap, with 70% achievable only in the highest quality materials. In addition, the tandem device photopotential needs to exceed the water-splitting potential in addition to overpotentials, typically requiring around 1.6 V. The combined bandgap of the two component semiconductor materials would then need to be somewhere between 2.3 and 3.2 eV, with the high end of the range for most materials, and the low end only for the highest quality materials. There are many pitfalls and extra losses in real-world multijunction implementations, so the general rule is no guarantee of successful water splitting.
It is important to emphasize that even if the tandem device develops sufficient potential to split water, the STH efficiency will always be restricted based on the lowest photocurrent in the stack, usually in the highest bandgap cell. As an illustration, TiO₂ with bandgaps of 3.1 eV could be combined with a low bandgap semiconductor such as silicon (1.1 eV) in a hybrid tandem to meet the photopotential requirement for PEC water splitting. As seen in Fig. 7.17, however, independent of the bottom junction, the device performance will never exceed approximately 2.5% STH. In this example, and in any series-connected tandem configuration, the performance is bound by the photocurrent limits of the highest bandgap PEC semiconductor. As an alternative hypothetical example for achieving STH efficiencies above 10%, a high-quality top cell semiconductor with a bandgap of 2.0 eV can be coupled with a 1.0 eV semiconductor for the bottom cell. This combination satisfies the combined bandgap requirement, and is within the 12–19% STH range, in fact, approaching 19% as seen in Fig. 7.23.

As discussed in Sect. 7.3.1, fundamental thermodynamics limits conversion efficiency for both single- and multijunction semiconductor PEC devices have also been established in numerous studies. The thermodynamic tandem limit under ideal conditions, i.e., including thermalization loss, but with no overpotential loss, has been calculated at approximately 40% STH. This is consistent with the optical limits shown in Fig. 7.23. It is also substantially higher than the 30% STH limit established for the single-junction case. Moreover, employing advanced multi-exciton device schemes which eliminate part of the thermalization loss, the theoretical tandem limit is further increased to 46% STH [58].

As also discussed in Sect. 7.3.1, more detailed analyses of multijunction PEC devices including optical, thermalization, and overpotential losses have been reported. Figure 7.24 shows the maximum achievable photocurrent density and possible associated STH conversion efficiency results from the Rocheleau and Miller analysis [82] for tandem devices of high-quality III–V semiconductors. This analysis, which focuses on devices operating in the high-efficiency region seen in the optical limits plot of Fig. 7.23, includes estimates for solid-state and electrochemical losses. As expected, the STH efficiency limits are somewhat lower than the optical limits, though they still do exceed 25% over a limited range of top and bottom cell bandgaps. The high-quality factor of the modeled semiconductors clearly plays a large role in maintaining high levels of performance. For lower quality thin-film materials, larger efficiency hits would be expected. Nevertheless, there should be adequate margins for achieving over 10% in thin film tandem devices under some circumstances [82].

As an illustration of this, the analysis of efficiency limits for triple junctions based on lower quality amorphous silicon materials were also reported by Rocheleau and Miller. In that analysis, a bottom cell, with a bandgap of 1.45 eV was assumed, and the middle and top cells were modeled based on the available range of bandgaps in thin-film silicon materials. The resulting current density and STH efficiency limits are shown in Fig. 7.25. Interestingly, even with the lower quality thin film materials, and despite the further photocurrent limits in the triple junction case, STH efficiencies in the 10–13% STH range appear readily attainable.
Fig. 7.24 Maximum achievable photocurrent densities and PEC STH efficiencies in tandem III-V semiconductor stacks as a function of top and bottom cell bandgaps based on analyses of Rocheleau and Miller [82] which included solid-state and electrochemical losses.

Fig. 7.25 Maximum achievable photocurrent densities and PEC STH efficiencies in triple junction amorphous silicon-based semiconductor devices as a function of top and middle cell bandgaps with a bottom cell gap of 1.45 eV. These results are based on the analyses of Rocheleau and Miller [82] which included solid-state and electrochemical losses.
Unfortunately, achieving high STH efficiency levels in laboratory prototypes of thin-film multijunction devices has been elusive to date. Challenges in identifying the most appropriate semiconductor materials to incorporate into a stack, and in developing sufficient solid–solid and solid–liquid interfaces have limited progress. Still, some encouraging research pathways using different device scheme approaches are ongoing.

7.6 PEC Multijunction Device Schemes

A number of different multijunction device schemes incorporating a combination of PEC cells, or a combination of PEC and PV cells have been explored in attempts to enhance solar hydrogen production via water splitting [9, 39, 105]. Some of the device approaches which have demonstrated varying levels of success to date include the following:

1. PV-electrolysis devices with multijunction solid-state PV cells coupled to electrolyzer systems.
3. Hybrid photoelectrode PV–PEC tandem devices comprising a PEC photoanode or photocathode integrated with a “buried” single-junction PV cell.
4. Hybrid photoelectrode PV–PV–PEC triple-junction devices comprising a PEC photoanode or photocathode integrated with a “buried” double junction PV cell.
5. Photocatalyst Z-scheme systems incorporating hydrogen-evolution and oxygen-evolution photocatalysts in single or dual bed configurations.

In the discussions of these approaches below, schematics of representative device configurations are included, accompanied by the basic device band structures to indicate the energetic driving forces behind the photolytic water splitting. Some examples of implementation in laboratory-scale prototypes are also presented.

7.6.1 Multijunction PV-Electrolysis

Establishing an interesting reference case for multijunction solar water splitting, multijunction PV cells with adequate useable photopotential to directly drive electrolysis have been coupled with electrolyzer systems, which can be separate or fully integrated. This approach is compatible with the “PV-electrolysis” reactor types described in Sect. 7.2.4. However, since the PV output of the multijunction cell is designed to have direct compatibility with electrolysis, no power conditioning units are required. A device level schematic is shown in Fig. 7.26 for the case of a triple-junction PV cell driving the electrolyzer reactions. The useable potential
developed by the triple-junction is capable of simultaneously driving photoholes at the electrolyzer anode surface and photoelectrons at the electrolyzer cathode surface. A unique set of design considerations apply in this case. The standard design strategies of current and lattice matching are necessary in the PV semiconductor stack. Additionally, the PV output voltage and current need to be sized to match the optimal operating conditions of the electrolyzer system. The electrolysis process can be independently optimized with appropriate electrodes and catalysts. Since there are no PEC cells in the device stack, any issues of semiconductor durability in solution are avoided. Auxiliary system losses including the lateral collection of current from the PV system and transmission of electrons through external wiring are inevitable, but can be minimized. Also inevitable, cost remains an issue in this approach. As discussed in Sect. 7.2.5, using commercialized PV and electrolyzer technologies, the hydrogen production cost would be greater than $10/kg, far exceeding the US DOE targets of $2–4/kg.

As a potentially lower cost implementation of the PV-electrolysis approach, triple junction amorphous silicon solar cells have been fully integrated with on-board electrolysis units. The amorphous silicon cells generate photovoltages over 2 V, and can convert sunlight to electricity with a stabilize PV efficiency of approximately 8–10%. At the same time, the electrolysis system has been
optimized to convert electricity to hydrogen at efficiencies near 70%. Net STH efficiencies from 5 to 7% are then achievable, as reported by Kelly and Gibson [106]. The approach is viable, but cost estimates have not been fully developed to evaluate how pragmatic it could be.

Higher efficiency, higher cost multijunction solar cells, such as the III–V materials systems capable of 20–40% PV conversion efficiency are also commercially available. Using these materials, 14–28% STH would be achievable using PV-electrolysis. In today’s market, however, such systems would be prohibitively expensive for any large-scale deployment. The 5–7% mark for lower cost amorphous silicon technology is a more appropriate near-term benchmark for practical solar hydrogen production. Alternative PEC-based schemes need to meet or exceed this benchmark to be viable.

7.6.2 PEC–PEC Tandem Photoelectrodes

PEC–PEC tandem photoelectrode systems for solar water splitting have been investigated as one alternative. These comprise separate photoanodes and photocathodes that are optically stacked in series and electrically interconnected, for example, using external wires. The two-photoelectrode approach is compatible with the photoelectrode-based reactor schemes, and subject to the associated reactor-level considerations described in Sect. 7.2.4. A device level schematic of the photoelectrodes is shown in Fig. 7.27, depicting a photoanode stacked in front of a photocathode. The electrode order can be reversed, however, if required by design consideration. For example if a higher bandgap photocathode material is available, this should be the top junction for optical reasons. As seen in the band diagram of Fig. 7.27, the back contacts of two electrodes are directly interconnected, and therefore their electronic levels align. Under illumination, the level of photoelectrons is driven “up” in the photocathode, and the level of photoholes is driven “down” in the photoanode, resulting in sufficient separation to split water.

There are a number of significant design challenges in this approach. The PEC–PEC arrangement requires two different stable PEC semiconductors, one n- and p-type, with added demands of optical compatibility for current matching, and band alignment compatibility for efficient photovoltage generation. As an extra consideration, the front junction must be fabricated on a transparent substrate coated with a transparent conductive layer to enable both light transmission and electron transfer through to the back junction. This adds complication, optical losses, electrical losses, as well as material cost in the photoelectrode fabrication. Additional conductivity losses can be associated with electrode spacing and gas separation. Due to these challenges, efficient PEC/PEC tandem systems have not been demonstrated to date. Proof of concept experiments have been reported, though. For example, a zinc-doped p-type Fe₂O₃ thin film photocathode has been used with a n-type Fe₂O₃ photoanode with reported efficiencies of 0.11%
An interesting recent innovation in this area has been the exploration of novel, futuristic tandem nanorod electrode structures using silicon and metal oxides [105].

7.6.3 PV–PEC Tandem Hybrid Photoelectrodes

An alternative tandem device being developed for solar water splitting is the PV/PEC hybrid photoelectrode. In this approach, a PEC top cell is monolithically stacked with a solid-state single-junction PV back cell. Such a photoelectrode device is compatible with both the one- and two-electrode reactor configurations described in Sect. 7.2.4, and the reactor-level considerations described in that section apply. Figure 7.28 shows a device level schematic for a specific two-electrode implementation incorporating a hybrid photocathode and a separate counter electrode. Possible alternative variations include the use of a hybrid photoanode instead of the photocathode, and the integration of the counter and
photoelectrodes into a single monolithic device. As seen in the band diagram of Fig. 7.28, the back contact of the PEC cell and the front contact of the PV cell are electronically interconnected via a tunnel junction. In addition, the back contact of the PV cell is wired to the counter electrode, and therefore electronically aligned. As illustrated in the band diagram, under illumination the photoelectron levels at the electrolyte interface and the photohole levels at the counter electrode are separated sufficiently to split water.

In contrast to the PEC–PEC tandem, the two junctions are stacked onto the same substrate, so no transparent substrate is needed. Also, in single-electrode implementations of this approach, electron transmission through wires is eliminated, reducing ohmic losses. There are some clear benefits, but also some specific design challenges associated with the PV–PEC tandem. Appropriately matched PV and PEC semiconductor materials systems are needed, which is challenging enough. In addition, process compatibility in fabricating the PEC cell directly onto the PV cell can be problematic; for example, the necessary restrictions on processing temperature can adversely affect device quality and efficiency. As an alternative strategy for avoiding the process compatibility issues, the PV and PEC cells can be

---

**Fig. 7.28** PEC–PV monolithically stacked hybrid water-splitting device
separately fabricated and mechanically stacked. This, however, adds manufacturing complexity and new losses. Similar to the tandem PEC–PEC case, stability of the PEC interface is a primary concern in the PV–PEC tandem. There are also similar conductivity losses associated with electrode spacing and gas separation. Overall, it is felt by many that the benefits can outweigh the challenges.

In evidence, the hybrid tandem approach has been under investigation for more than a decade [1, 73, 108, 109]. The best laboratory-scale demonstration to date is unquestionably the NREL’s III–V tandem hybrid photoelectrode structure, which still holds the world-record STH efficiency for PEC water-splitting [1, 2]. This hybrid device, shown in Fig. 7.29, consists of a p-type GaInP$_2$ PEC electrode interconnected by a tunnel junction to a buried GaAs p/n PV cell grown on a GaAs wafer substrate. The basic band structure is equivalent to the one shown in Fig. 7.28. In the NREL device, the top junction GaInP$_2$ bandgap is 1.83 eV and the back junction GaAs bandgap is 1.44 eV. Under illumination, hydrogen is evolved at the front photocathode surface and oxygen is evolved at the counter electrode. Under concentrated 11 sun illumination, and in 3 M sulfuric acid electrolyte, prototypes of this device have demonstrated over 12% STH efficiencies for up to 20 h. Subsequent drops in photocurrent and efficiency were reported due to corrosion of the GaInP$_2$. Circumventing the stability limitation, the basic PV structure of this device has also been integrated in tandem-PV-electrolysis prototypes, demonstrating STH efficiencies exceeding 16%. In attempts to enhance the tandem PV–PEC implementation, fundamental research to stabilize the GaInP$_2$ PEC interface is ongoing [97]. Though lacking long-term stability, and reliant on high-cost III–V materials, the results of the NREL demonstrations affirm the potential for high efficiency in multijunction PEC systems. As a “sanity check” for the high-performance levels

![Fig. 7.29 Device schematic of the NREL GaAs/GaInP$_2$ hybrid tandem device submersed in the electrolyte. Light enters through the top GaInP$_2$ PEC electrode then filters down to the GaAs solar cell. This device has demonstrated spontaneous water splitting under illumination with over 12 STH%](image)
achieved, it is interesting to note that, according to Fig. 7.23, the STH efficiency limit for a III–V 1.83 eV/1.44 eV tandem stack, is about 18% STH, consistent with the experimental results.

### 7.6.4 PV–PV–PEC Triple Junction Hybrid Photoelectrodes

A further extension of the PV–PEC hybrid photoelectrode device approach is the PV–PV–PEC triple-junction hybrid photoelectrode. The triple-junction variation is capable of developing higher photopotential, but at the cost of reduced photocurrent and added device complexity. In this configuration, a PEC top cell is monolithically stacked on a solid-state double-junction PV cell. Again, the resulting device is compatible with both the one- and two-photoelectrode configurations in reactor types described in Sect. 7.2.4. A specific two-electrode implementation with a PV–PV–PEC hybrid photoanode and a separate counter electrode is schematically illustrated in Fig. 7.30. Of course, alternative variations include the use of a hybrid...
photocathode instead of the photoanode, and the integration of the counter electrode into a single electrode device. Similar to the PV–PEC tandem case, the back contact of the PEC cell and the front contact of the PV–PV double cell are electronically interconnected. Also, the back contact of the PV–PV cell is electronically aligned with the counter electrode. As seen in the band diagram, under illumination the photoelectron levels at the electrolyte interface and the photohole levels at the counter electrode are separated sufficiently to split water. Compared with the PV–PEC tandem, additional photopotential is generated in the PV–PV buried cell of this triple-junction device, offering compatibility with a broader range of lower performing PEC materials. The same design considerations described for the PV–PEC tandem hold, though the extra PV cell adds further complexity and cost in terms of both materials and manufacturing.

The triple-junction hybrid approach has also been well studied, particularly in applications utilizing thin-film metal oxide photoanodes, which frequently require high levels of internal biasing to compensate for unfavorable band positioning at the electrolyte interface. Tungsten trioxide (WO₃), for example, has a proven track record in the stable photo-oxidation of water. However, due to its band structure, biasing in excess of 1.0 V has been needed to achieve water splitting [110]. This has necessitated using a double-junction PV buried cell in triple-junction hybrid photoelectrode applications. At the University of Hawai‘i (UH) at Manoa, n-type WO₃ thin-film electrodes have been mechanically stacked on top of amorphous silicon (a-Si) tandem PV cells into a PV–PV–PEC hybrid device, as illustrated in Fig. 7.31a [7]. The “load-line” analysis of this device [54] is shown in Fig. 7.31b, superimposing the independently measured two-electrode response curves of both the WO₃ PEC (in 0.33 M H₃PO₄) cell and the a-Si tandem PV cell under AM1.5G illumination [57]. Consistent with the operating point indicated by the intersection in the load-line plot, the device successfully split water, with a stabilized photocurrent of 2.5 mA/cm², corresponding to a STH conversion efficiency of approximately 3.1%. An interesting and important insight from the load-line plot is that the limiting factors to high efficiency in this specific device are the WO₃ onset potential, saturation photocurrent, and fill factor. If these could be adequately addressed, such a triple-junction approach using low cost thin film semiconductor materials has clear potential for exceeding 5% STH. Based on the optical limits, as seen in Fig. 7.25, successful bandgap reduction in the WO₃ would be needed to exceed the 5% mark. Alternatively, new lower bandgap thin-film materials may emerge through research for this application.

In addition to the amorphous silicon implementation, several interesting variations of the PV–PV–PEC hybrid device based on WO₃ photo-oxidation cells have been demonstrated focusing on low-cost material components. Most notably is the impressive work using tungsten oxide PEC cells in conjunction with dye-sensitized solar cells. The Grätzel–Augustynski “Tandem Cell” is the best example of this [5], though some more recent work has been reported by the Arakawa group in Japan [111]. Efficiencies in these all-PEC configurations are reported in the 2.5–4.5% STH range. Moreover, there has been commercial interest in this approach.
Another interesting example has been the development of an all thin-film silicon triple-junction hybrid device by MVSystems, Incorporated [109, 112–114]. In this device, hydrogenated amorphous silicon-carbide alloy thin-films are used for all three junctions, including the PEC junction and both buried PV cells. Bandgap tuning through alloy composition variations is used for current matching. Compatible with the triple-junction efficiency analysis shown in Fig. 7.25, such a device has theoretical potential for exceeding 10% STH. Laboratory prototypes, though, have been limited by high-loss potential barriers formed at the PEC interface, and demonstrations to date of spontaneous solar water splitting have performed closer to 1% STH. Ongoing work to reduce the interfacial barrier issues offers much promise for achieving commercialized PEC hydrogen production based on thin-film silicon alloys.

Fig. 7.31 (a) Device schematic of the UH WO₃/a-Si/a-Si hybrid triple junction device submerged in the electrolyte. Light enters through the top WO₃ PEC electrode then filters down to a-Si/a-Si tandem solar cell. This device has demonstrated spontaneous water splitting under illumination with over 3.0 STH%. (b) Load-line analysis of the device, showing operating point of 2.5 mA/cm² based on superposition of the PV and PEC responses.
7.6.5 Photocatalyst Z-Schemes

The “Z-scheme” approach, which is receiving much recent attention, particularly in the Japanese research community, is essentially a photocatalyst version of the tandem PEC–PEC photoelectrode system. As shown in the schematic representation in Fig. 7.32, n-type photocatalyst particles drive the hydrogen evolution reaction (HER) and p-type photocatalyst particles drive the oxygen evolution reaction (OER), achieving water splitting in a tandem fashion. In the Z-scheme, electrons are exchanged between the photoanode and photocathode particles via an ion mediator in solution (such as the Fe$^{3+}$/Fe$^{2+}$ couple), rather than through wires, as in the PEC–PEC photoelectrode case. Slurries of the two photocatalysts can be mixed together in single-bed reactors, or alternatively, slurries of the HER and OER photocatalysts can be separated into two beds, connected with an ionic bridge, consistent with the one- and two-bed reactors discussed in Sect. 7.2.4.

As indicated in the DTI techno-economic analysis report, the particle-based systems offer some of the best hope for low-cost hydrogen production on a large scale, but there are a number of unique design challenges posed by this approach. For example, charge separation within each catalyst particle needs to be enhanced. The photocathode particle has to simultaneously reduce hydrogen and oxidize the mediator at separate surface sites. Similarly, the photoanode particle must both
oxidize water and reduce the mediator. Functionalization of surface sites, for example, with reaction-specific nanoparticle catalysts, is necessary to draw apart photogenerated electrons and holes, and to enable the simultaneous reactions.

As a particular challenge in the single bed reactor system using Z-scheme photocatalysts, both hydrogen and oxygen are evolved in the same reactor bed, with associated safety hazards. Safe gas separations techniques could be employed, but at substantial additional cost. The dual bed system avoids any possible explosive hydrogen/oxygen gas mixtures, but the greater land area requirements and conductivity losses in the ion bridge reduce the STH performance. Both the single- and dual-bed systems have their advantages and disadvantages, but either way, new HER and OER photocatalysts need to be identified with sufficient efficiency and stability for practical hydrogen production.

The search for effective HER and OER photocatalyst materials for Z-scheme systems is ongoing. Extensive libraries of multicomponent metal oxides and oxinitrides are being developed and screened [9, 96–99]. Specific material systems have demonstrated spontaneous water-splitting under visible light, for example, SrTiO3:Rh combined with BiVO4 in the presence of an Fe3+/Fe2+ redox couple [115], but STH efficiencies remain extremely low. Visible light water splitting has also been accomplished using a RuO2–TaON and Pt–TaON mixture with I3–/I− redox couple, but again with low STH efficiency [116]. A recurring theme in PEC research is that new materials continue to offer new hope.

7.6.6 Many Options

Compared with the photovoltaics industry, the field of multijunction PEC hydrogen production is relatively new, but there has been substantial exploratory research in recent years. In addition to the above examples in PV-electrolysis, multijunction photoelectrolysis, and Z-scheme photocatalysis, numerous variations which mix and match concepts and components of all three are conceivable. In the final analysis, commercialization will depend on balances between performance, complexity, and cost. The simpler the system, the better. Single-junction photocatalyst or photoelectrode systems offer the greatest simplicity, but the significant difficulties in achieving high efficiencies in these systems necessitate a serious look at the multijunction alternatives. Efficient tandem systems are the next best hope. Analysis has validated the potential for high STH conversion efficiencies in tandem double- and triple-junction schemes and there has been experimental verification, though still limited, on the laboratory scale. Real challenges remain in the identification of low-cost material systems capable of reproducing the laboratory performance levels on a commercial scale. As the research community advances in its ability to engineer high-efficiency absorber materials and functionalized interfaces, and as it learns to integrate these in optimal multijunction designs, the opportunities for success will continue to grow.
7.7 In Summary

There have been significant R&D efforts over the past several decades advancing multijunction device technologies, particularly in the PV field. These have resulted in the highest solar-to-electric conversion efficiencies to date, exceeding 40% in high quality III–V cells being commercialized for space applications. These ultrahighly efficient cells are, however, expensive, and not practical for large-scale deployment in solar power production plants. The multijunction approach continues to make strides in improving performance in lower cost thin-film solar cells, including, for example, amorphous silicon and copper chalcopyrite devices, though the commercial benefits have not yet been fully realized. It is nevertheless clear that low cost and high performance are the keys to the success of transformative scale solar electricity. Multijunction thin film devices offer hope in this area, and this hope extends to PEC hydrogen production.

There are more recent efforts to develop multijunction devices and systems for solar hydrogen production, though this field is still in its infancy in comparison with PV technologies. STH conversion efficiencies of 12–16% have been demonstrated on the laboratory scale in tandem devices using high-quality III–V materials, but with limited durability. Multijunction PEC and PV-electrolysis systems using lower cost thin-film materials have demonstrated better stability, but at conversion efficiencies limited in the 3–8% range. Ultimately, durable, low cost systems with STH conversion efficiencies over 10% are needed for commercial deployment of solar hydrogen technologies. Fortunately, continued development of multijunction schemes offers great promise, specifically in terms of enhanced photopotentials and broadened spectral bandwidth. Analyses of tandem and triple-junction PEC schemes indicate that efficiencies exceeding 10% STH and in special cases up to 25% STH, are in fact possible. Technical barriers have stood in the way of achieving such performance levels, but these are solvable if the fundamental problems are broken down and appropriately addressed.

7.7.1 What Is Needed?

Based on PEC research efforts to date, it has become increasingly clear that new, innovative material systems, interfaces, and devices are needed to achieve high-performance, low-cost solar water splitting. Breaking this down into specifics relative to the multijunction approach:

- Low-cost semiconductor materials with appropriate bandgap and sufficient opto-electronic quality need to be identified and optimized to enable efficient solar energy absorption and charge extraction.
- Effective PEC interfaces activated for the gas evolution reactions and passivated for corrosion and other parasitic reactions need to be engineered, using, for example, heterojunction and/or catalytic coating and/or dispersions.
Low loss solid-state interfaces need to be engineered for bridging component cells in a multijunction device structure, for example, using tunnel junctions or intermediate transparent conductive oxides layers.

Innovative multijunction designs combining the best materials and interfacing schemes into efficient photoelectrode or photocatalyst device structures will be the key.

Innovative synthesis and fabrication techniques will be needed to enable cost-effective large-scale manufacture of multijunction photoelectrodes or photocatalysts.

Leveraging the knowledge gained from the PV community in the successful development of high-efficiency multijunction device designs will be integral to success. Key advances, however, are needed specifically in the development of the PEC material components and interfaces, since these remain the greatest technical barriers.

### 7.7.2 What Is Happening

In efforts to address the technical barriers, the PEC research community over the past several years has been developing new theoretical, synthesis and characterization tools using improved state-of-the-art methodologies [57]. These tools are being employed in several parallel strategies for developing the innovative materials and interfaces necessary to the success of PEC hydrogen production systems.

One specific approach is the further development of the traditional PEC semiconductor materials in thin-film and nanostructured forms for higher efficiencies. There have been significant scientific efforts in this approach, including continued development of the following materials classes:

- Tungsten-oxide and related bandgap-reduced compounds [6, 117–121].
- Iron-oxide and related enhanced mobility compounds [122–126].
- Vanadates, including bismuth vanadate and related compounds [127–129].
- Titanium dioxide and related bandgap-reduced compounds [61, 130].

This list is by no means comprehensive, but does represent a large part of the research endeavors to advance well-established PEC semiconductors.

Another important strategy is the adaptation of efficient PV semiconductor thin-films and nanostructures for effective use in PEC applications. Recent research in this area has focused on material classes with inherent bandgap tuning capabilities such as the amorphous silicon compounds (including silicon carbides and nitrides) [109, 112–114, 131, 132], and polycrystalline copper chalcopyrite compounds [133–137].

Some of the most recent innovative approaches have included the development of entirely new materials classes and structures for use in photoelectrode or photocatalyst devices. Examples include quantum-confined WS$_2$ and MoS$_2$...
nanoparticle photocatalysts embedded in a conductive scaffolding [138] and silicon nanorods embedded in membrane structures [105, 139]. Also cutting-edge are efforts in the development of breakthrough synthesis technologies to reduce the cost and enhance the stability of the high-performance III–V multijunction structures that have demonstrated the best efficiencies to date, both in PV and PEC [4, 87]. Future progress in all these approaches, which will be integral to the ultimate success of PEC hydrogen production, will depend on the best use of the most advanced research tools.

### 7.7.3 The Path Forward

For all PEC systems, and for multijunction water-splitting systems in particular, the fundamental solutions reside in MATERIALS and INTERFACES. Good solid-state light absorber materials with efficient charge transport properties that can be synthesized inexpensively are the cornerstone. Low loss interfaces are the next major step. For example, low-defect and low-recombination solid-state interfaces are essential between cells in a multijunction stack, and at back contacts. These are achievable, as demonstrated through years of R&D in the PV research community. More challenging are the solid–liquid electrochemical interfaces in PEC multijunction schemes. Functionalizing surfaces to produce the right energetic and kinetics for promoting the water-splitting half reactions while inhibiting parasitic reactions and recombination loss will be the key to unlock the full potential of PEC hydrogen production.

The lessons learned to date in the PEC research and development community, coupled with PV knowledge and experience amassed over decades could be extremely fruitful, both for solar hydrogen and solar electricity production. Some questions to consider for the pathways forward include the following:

- Can functions of light absorption, charge extraction, reaction catalysis, and production management be effectively separated, and independently optimized in PEC devices? If so, it opens up the design space considerably. For example, a good absorber material would not need to be catalytic, but could be coupled with a good catalyst material with the design of an efficient interface.
- Can cost-effective thin film PV materials (e.g., amorphous silicon or more efficient copper chalcopyrites) and multijunction devices be adapted through surface modifications to split water? If so, a wealth of PV manufacturing experience would be available for enabling commercial scale deployment.
- Can new low-cost synthesis routes be devised for ultrahigh efficiency III–V materials and devices? If so, the barriers to large-scale deployment of solar electricity are largely gone. Moreover, if appropriate surface modifying materials and interfaces are also developed, then the commercial viability of PEC hydrogen production would be assured.
Can new forms of earth-abundant materials, including oxides, nitrides, and sulfides, be discovered and developed with sufficiently good light absorbing and charge transport properties for solar energy conversion? This, with the development of efficient solid-state and electrochemical interfaces, and with the design of effective multijunction photoelectrode or photocatalyst systems, could revolutionize both PV and PEC technologies!

These questions encompass a broad vision and provide a coarse roadmap for paths forward. The pathways to success of PEC hydrogen production are along the road of technical collaboration; and the trail will be blazed by the new generation of scientists and engineers.

Acknowledgments

The authors would like to acknowledge and express their great admiration for all members of the international PEC research and development community; with special nods of appreciation to the US Department of Energy’s PEC Working Group supported by the Fuel Cells Technologies Office, and to Annex-26 of the International Energy Agency’s Hydrogen Implementing Agreement. They also thank the members of the Thin Films Laboratory at the University of Hawaii at Manoa’s Hawaii Natural Energy Institute, including Drs. Nicolas Gaillard, Bor Yann Liaw, Yuancheng Chang, and Richard Rocheleau, as well as Jess Kaneshiro, Jeremy Kowalczyk, Xi Song, and Brett Ikei for their encouragement and support to this effort.

References


135. Bär, M., Böhne, W., Röhrich, J., Strub, E., Lindner, S., Lux-Steiner, M.C., Fischer, Ch-H: Determination of the band gap depth profile of the penternary Cu(In$_{1-x}$Ga$_x$(S$_{1-y}$Se$_y$)$_2$ chalcopyrite from its composition gradient. J. Appl. Phys. 96, 3857 (2004)


Part IV
Future Perspectives
Chapter 8
Economic and Business Perspectives

Julian Keable and Brian Holcroft

8.1 Introduction

Mankind faces the biggest challenge in the history of energy use – to find how to store natural energy in a way that would replace the vast store created by great nature herself, a store of which we have used a half in the last 50 years [1].

This is also one of the biggest opportunities ever to arise. The market is vast for a fuel that can be conveniently stored and transported, provided from a sustainable, low environmental impact source, at an affordable price, to meet the inevitable shortfall after “peak oil.”

There are many possible ways forward. One generic way is to split water, and store the resulting hydrogen. Photoelectrochemical (PEC) hydrogen production is not yet a commercial solution for this, and for it to be considered there are basic material improvements required, together with robust system implementation. Materials with parameters in the range of photocurrents of 8 mA/cm² in “1 Sun” with 0.8 V anode to cathode bias would open the way to exploitation of PEC-derived hydrogen, and the possibility of a cost/m² under $80, excluding PV bias costs with lifetimes >10 years. An early demonstration model of such a system is shown in Fig. 8.1.

8.2 The Need

Our entire world economy has grown out of the availability of abundant and cheap energy supplies. The cost of energy is now under threat, and the future is by no means secure.
8.2.1 Increased Demand/Reduced Supply = Rising Cost

In the 4 years up to 2007 [2],

- Oil consumption increased by 3 million barrels/day
- Natural gas increased by 8 billion cubic feet/year
- Coal by 1.4 thousand short tons/year
- Electricity by 2.1 billion kWh/year

Projections show that demand continues to grow – and that the transport sector is elastic to price, there being no available alternative.

Oil production per capita fell from 5.26 barrels in 1980 to 4.73 barrels in 2006. The DOE/EIA give a price projection from [2]:

- $60/barrel in 2009 to
- $110 in 2015
- $115 in 2020
- $130 in 2030

Fig. 8.1 Demonstration model of a 1-m² PEC array
8.2.2 Growth in Energy Demand Arises from Several Different Causes

These causes are

- Population growth – mainly in the developing world
- Increased standards of living within upper income sectors
- Increased size of middle income groups

The figures quoted above hint at the size of this growth in energy demand. The EIA forecast an 80% increase in consumption in non-OECD Asia, largely due to transport needs [3]. Of special interest is the ever increasing proportion of energy use in the transport sector generally, in particular road vehicles and air travel. Of total primary energy use, some 40% is used in the transport sector. This is met by the use of oil in different grades (petrol, diesel, jet fuel), and there is little prospect of a change to other energy sources here. There is an absolute need for transport fuel to be convenient and transportable – on board – which means that the use of electricity is itself dependent on storage, in batteries of some sort.

8.2.3 “Peak Oil”

Peak oil was first predicted by M. King Hubbert in 1956 [4]. He showed that every well increased its production up to a peak, coinciding with the depletion of half the total capacity of the well, then slowly declined once past this peak. He showed that this is just as true for the overall capacity of all wells. Dr. C.J. Campbell, a trustee of the Oil Depletion Analysis Centre in London, spelt out the consequences in some detail in his 1997 book “The coming Oil Crisis” [5]. In 1980, oil production first surpassed new oil discoveries. Therefore, there is a need for an alternative to augment oil and preferably from a low impact source due to climate change issues.

Renewable sources are attractive for sustainability – e.g., wind, hydro, tidal, solar etc., and geographic factors will determine where each is appropriate. However, the sheer size of overall demand points to solar – the solar resource far outstrips all others. With a global usable energy flux of 120 PW, it is some 8,000 times the current total energy use, and able to supply growing demand as far ahead as we can see [6].

There has been a huge emphasis on the production of electricity from alternative sources, and for many, “solar cells” mean photovoltaic cells. Wind, hydro, tidal, and atomic sources all tend to equate to the supply of electricity. As noted above, electricity needs to be stored if it is to meet the demands of transport, and, in the case of wind and solar, periods when the source is unavailable.
8.3 The Commercial Challenges

Energy must be affordable, but the cost of energy in different forms (with and without tax) varies widely.

8.3.1 Price Distortions

Governments distort true prices by giving subsidies on the one hand, and applying taxes on the other. The 2005 Energy Bill in the USA gave $6 billion for oil and gas, $9 billion for coal, and $12 billion for nuclear [7]. President Obama is seeking to remove those subsidies. Tax on petrol varies by country. In UK 67% is duty, in Russia 0%. In the USA, tax varies by state, with Alaska at 18 cents/gallon, and New York at 61 cents [8].

8.3.2 The Climate Change Factor

Despite the growing awareness and acceptance that climate change is real, and that it poses huge threats to our whole existence, there has been little appetite by Governments to address the issue in a robust way. The conclusion drawn from this may be that carbon-free alternatives must compete without relying on government support or coercion.

8.3.3 The Need for Clear Focus

Research effort, and the subsequent process of bringing research results to a successful product and market, is costly. It is important, therefore, to focus on well-formulated objectives. Set out below are a number of such objectives, all having in view the aim of finding a viable way to store solar energy, as an alternative to mining the natural resources we have inherited.

8.3.4 Definition of Aims

Development of an energy supply which is

- As cheap as existing “conventional” supplies
- As convenient in use
- Less environmentally damaging
- Reasonably safe in use (all energy can be dangerous!)
8.4 Storage of Solar Energy: The New Challenge, and The Hydrogen Solution

Throughout history, mankind has existed by harvesting and mining natural resources (wood, coal, oil, and gas) for his energy needs. Only with the coming of the atomic age can it be said we have ourselves invented or created energy, and even now we have still to find anything comparable with the stored energy created by nature. Her “invention” of the tree, which covers itself with disposable solar collectors during the sunny season, converts sunlight to a form of stored energy, then discards those leaves when winter comes – how elegant!

Wood, coal, oil, and gas can and do store energy for any amount of time, ready for instant use.

This is our real challenge, for whether “peak oil” has already come, or whether it is still 20 years away (“tomorrow morning”), once we are on the far side of the mining curve nothing can replace that golden resource.

So we come to our subject here – the possibility of splitting water using sunlight, to provide a source of stored energy at an affordable price. PEC is such a possibility, but many hurdles remain.

8.4.1 Can We Learn from Nature?

The sunlight/photosynthesis/fossil fuel process has been working at efficiencies well below 1%, thus needing vast land areas – land areas now committed to other uses where they are fertile, or lacking either water or sun elsewhere. There are many ways to approach this problem, PEC being only one, and this is not the place to examine the alternatives. Each has advantages, disadvantages, risks, and state of readiness.

8.4.2 There Are A Number of Ways to Split Water and Yield Hydrogen

Direct thermal splitting is one such way – and there is work going on looking at efficient catalysts for this [9, 10]. Electrolysis is another and well-understood way – the use of electric power to split water. It may be driven by electricity from any source – conventional (fossil fuelled), wind power, hydro or wave power, nuclear or PV [11]. Steam methane reformation (SMR) is at present the cheapest way to produce hydrogen, and relies on fossil fuel (natural gas). This provides a cost per kg and kWh of H2 burnt split between capital and electrical power cost [12, 13].

The DOE target for hydrogen production from natural gas (NG) was $3/kg H2 by 2005, and $1.50/kg by 2010 [2]. The volatile price of NG is the main factor here.
If NG is traded at $1/kg equivalent, and if conversion from NG to H₂ is at a ratio of 1–2.4 (when carbon capture and storage is included) then even at very large scale the cost of H₂ is likely to be some $2.4/kg. Thus the $3/kg target, not yet reached, is likely to remain the lowest.

Although meeting a target cost of twice this figure using PEC is already very challenging, this aim must be kept in sight, and achievements measured in the light of such goals.

8.4.3 End Use of Hydrogen

The end use of hydrogen is, ideally, by direct combustion. Process heat, cooling by the absorption cycle, combustion in IC vehicle engines, cooking, and water heating all make the best use of hydrogen. To generate electricity, while perfectly possible, is to lose efficiency overall. Fuel cells have yet to be proven, and have yet to show cost benefits. Hydrogen has an important role to play without the need for deployment in that way.

8.4.4 EU/US Targets for H₂ Evolution

The US DOE target has been shifting gradually upward, from $1.50/kg delivered H₂ (2001) to $2/kg at production site, to $3/kg. EU targets twice the untaxed petrol cost, on the basis that petrol taxation will make H₂ competitive if untaxed.

8.5 PEC Issues

Each element of the PEC process has a bearing on the final cost of hydrogen production, and hence of commercial viability.

8.5.1 PEC Costs for Hydrogen

It is constructive to consider the optimistic cost of hydrogen produced today using PEC compared with the target prices. In this calculation, typical values of photocurrent for modified iron oxide of 2 mA/cm² at 1 “Sun” illumination with a bias of 1.2 Volts have been used and costs for 17% efficient PV at $2/W peak incorporated. Clearly, the cost of hydrogen per kg produced is dependent on the location and a site in Southern Spain has been assumed with an averaged monthly insolation of 4.8 kWh/m²/day. To produce 1 kg of hydrogen per day would require some
50 m² of PV and 300 m² of PEC with minimal overall losses and inefficiencies. Assuming that the PEC costs are around $200/m² (PV costs are $300/m²) then this yields a cost per kg of hydrogen of $20 (assuming a 10-year lifetime) for the capital costs alone. Even if the PEC costs were halved to $100/m² then the hydrogen costs are still $12/kg. At this lower price the PV costs are around half the PEC costs due to the large area of PEC required.

This cost of $20/kg is clearly too high and around a factor of 10 greater than the target price and does not include additional costs of land, maintenance, installation, transport, the balance of plant costs, and profit. Clearly, an improvement in PEC efficiency would be advantageous to limit the land usage. This suggests that significant improvements in performance of the PEC device are required to meet commercial demands.

We should not ignore competitive means of producing hydrogen from a solar input via electrolysis. Given the likely timescale for improvements in PEC, we should use a comparison with other methods under development. Direct comparison with low cost PEM electrolyzers powered by PV is instructive. The US targets for low cost PEM are $400/kW with 67% efficiency by 2012 and $125/kW with 74% efficiency by 2017. The reported status in 2007 was $987/kW with a cell efficiency of 67% [14]. Using this data for the site in Southern Spain with PV costs of $2/W peak we can estimate hydrogen costs for the capital alone with a 10-year payback if the systems were implemented of $8/kg using the 2007 figures, reducing to around $5/kg of hydrogen with the 2017 figures. The reduction in cost is not as marked as expected from the reduction in PEM electrolyzer costs due to the influence of the PV costs which have been assumed not to reduce below $2 per watt peak over the period. The PV costs would be 94% of the capital costs in such a system, and using the 2017 figures it would suggest this system would require around 60 m² of 17% efficient PV.

Using the 2012 figures the PV accounts for 80% of the cost of a PEM system, with the PEM electrolyzer capable of producing 1 kg of hydrogen per day in Spain. The estimated capital cost of some $4,000 yields a cost per kg of hydrogen of $6 assuming a 10-year payback.

8.5.2 Efficiency and Improved Photocurrent

It is widely held that the successful application of PEC to the water-splitting process must reach or exceed an efficiency of 10% (i.e., for a given incident solar energy, the energy content of the liberated hydrogen must be 10% or more). Put simply, at “1 Sun” (1,000 W/m²), the aim must be for 1 m² of device to liberate hydrogen having an energy content of 100 W. In the following sections, we set out to expand on materials parameters that support this.

The efficiency of PEC is directly related to the magnitude of the photocurrent produced by exposure to light. Clearly, doubling the efficiency or photocurrent halves the area of PEC required to produce a given mass of gas.
The solid line in Fig. 8.2 represents the capital cost per kg of hydrogen produced with a 10-year payback assuming a PEC device requiring a bias of 1.2 V that can be manufactured for $100/m². To match the low-cost PEM equivalent price of $6/kg would require an improvement in efficiency to yield a “1 Sun” photocurrent in excess of 8 mA/cm². At this value the area of PEC required in Spain would be around 75 m² compared to 50 m² of PV and 60 m² of PV for a low-cost PEM. As can be seen increasing the photocurrent further provides decreasing gains in lower cost hydrogen as the PV costs, which do not decrease with increasing PEC photocurrent, begin to dominate the total costs. At a “1 Sun” photocurrent of 12 mA/cm² the PV costs, assuming $2/W peak, would be $15,000 compared to $5,000 for the PEC. The most advantageous method to reduce cost further is to reduce the bias requirements and so reduce the amount of PV required.

### 8.5.3 Voltage Issues and Reduced Bias

Increasing the efficiency reduces the cost of PEC to produce a given mass of gas but not the cost of PV. Since PV dominates the cost overall, then lower cost of PEC does not itself produce a significant saving. How can we reduce the amount of PV power needed? This can only be achieved by reducing the bias voltage as the mass of hydrogen produced is proportional to the current.

Water splitting by electrolysis has a threshold minimum of 1.23 V and in a PEC device this is effectively reduced by the photocatalyst, the reduction being in part related to the band gap of the material and the positioning of the energy bands relative to the hydrogen and oxygen evolution potentials. However, an overvoltage is observed for realistic devices. The size of this overvoltage is critical and must be minimized.

By comparison of the solid line for a PEC device manufactured for $100/m² with a bias of 1.2 V and the single dashed line for a PEC device with a bias of 0.8 V (Fig. 8.2) it can be seen that the hydrogen cost per kg can be reduced below that of a

---

**Fig. 8.2** Capital cost of hydrogen per kg as a function of photocurrent
low-cost PEM at reasonable photocurrent values above 5 mA/cm$^2$. Efforts to reduce the bias are essential to achieve a meaningful advantage over low-cost PEM and ideally PEC systems requiring no external bias would be highly attractive.

It is important to note that with devices exhibiting “1 Sun” photocurrents of 8 mA/cm$^2$, currents of the order of 80 A for a single m$^2$ can be expected. A 1 mΩ resistance in a pathway carrying 80 A would result in a resistive voltage drop of 0.08 V which is significant when efforts to reduce the bias are hard won. Hence, very low electrode resistances and interconnections are crucial to a PEC system. Also sources of electrode overpotential, losses due to gas separation membranes, and reactor designs all play a part in minimizing the external bias requirements for a PEC device.

From a system viewpoint anything that loses voltage, e.g., electrode resistance, interconnections, over potentials, reactor design, electrolyte, H$_2$ evolving electrode, etc., is important.

### 8.5.4 Scale-Up Problems

Research is typically carried out using small samples on which to try alternative coatings or application processes. By this means, many great experiments can be made with a relatively small use of materials and other resources. However, in the context of PEC development, the question of electrical resistances must not be overlooked.

At small scale, only small travel distances will be involved – the question of resistance may not well arise. But once a promising process comes to be scaled up to commercially realistic units, current flows then rise from milliamps to amps, and the need for a low resistance pathway may rule out the concept of a semitransparent conducting coating on glass (for example) – only highly conductive metallic electrodes which are nontransparent may suffice.

### 8.5.5 “Real Estate” Issues

There are many ways in which the area of land required per unit of energy harvested can affect cost. These include location (level of insolation to be expected), urban or off-grid siting, the need for support infrastructure, and other competing land uses.

For example, if we compare the 2012 low-cost PEM figures with the PEC estimates based on performance today we see that the PEC system would require around 20% less area of PV and hence the cost of PV. However, the land area used by the PEC system today would be dominated by the 300 m$^2$ of PEC which is some five times the area of a compact PEM system in which the PEM electrolyzer does not need exposure to light and can be located beneath the PV. The capital cost of hydrogen produced per kg is around double from the PEC system assuming it can
be manufactured at costs of $100/m². This would suggest that to be competitive with proposed low-cost PEM it is essential to reduce the area of PEC required through increased efficiency.

### 8.5.6 Solar Spectrum Considerations

In general, it is clear that having PV and PEC side by side is less efficient than using the same area to harvest different parts of the spectrum for the two requirements, so light harvesting is itself an issue relating to cost.

The tandem/hybrid approach seeks to address this question. A collaboration between EPFL and the University of Geneva leads to a proposal for a “Tandem Cell” which sought to use different parts of the spectrum of white light for different purposes. A patent was granted for this idea, based on harvesting energy from the blue end of the spectrum in a PEC application, and the red end of the spectrum for generating the required bias to achieve water splitting, using PV.

This elegant solution appeared to offer a way of “taking two bites from the same cherry,” and did function at very small scale. Care with such designs must be taken not to limit the transmitted light through the PEC device and into the underlying PV, otherwise PV outputs will be reduced which will result in increased costs of the PV. The need for highly conductive substrates at any reasonable photocurrent density and area of device within PEC devices will severely limit the light transmission and limit the commercial viability of this approach.

### 8.5.7 Durability

Products must be durable, and pay-back calculations will be influenced by the expected product life-time. As can be seen in Fig. 8.2, even with an external bias of 0.8 V and a PEC manufacturing cost of $50/m², which is one sixth that of current PV and a “1 Sun” photocurrent of 12 mA/cm² the hydrogen cost of the capital costs alone are still in excess of the EU and US targets assuming a 10-year pay back. These targets can only be met by increasing the durability of the PEC to give lifetimes approaching 20 years. To put this in perspective a 20-year lifetime operating 9 h per day is in excess of 65,000 h, which is challenging when you consider that a 10,000-h auto engine test would be considered impressive.

Although there may well be markets for shorter life-time products, in the longer term there will be an absolute requirement for products capable of operating for 15 or 20 years, comparable with other core technologies.

For these reasons, durability testing should be built into the overall PEC development program from the earliest stages. In some cases, it will be found that durability cannot be achieved for reasons of fundamental material science, thus in such cases there will be little reason to pursue an otherwise interesting line.
8.5.8 Commercially Attractive Targets

Figure 8.3 highlights that for PEC with a 1 Sun photocurrent of 8 mA/cm² and a bias of 0.8 V, capital costs of hydrogen per kg less than $3/kg, with a 10-year payback, cannot be achieved (solid line). Equally, PEC at similar photocurrents with 1.2 V bias cannot achieve sub-$3/kg hydrogen costs even with a 15-year payback (large dashed line). To achieve sub-$3/kg hydrogen capital costs for PEC with photocurrents of 8 mA/cm², the bias must be at least as low as 0.8 V with a 15-year payback and PEC costs below $100/m² (small dashed line). If the bias can be reduced further to 0.4 V then with a 15-year payback PEC costs as high as $150/m² can be sustained, which is around half the current cost of PV.

Clearly, the performance of PEC today is not ready for commercial exploitation. However, with defined improvements it offers a route to the production of hydrogen on a large scale which would meet EU and US targets and be comparable with current methods of hydrogen production. To achieve this, “1 Sun” photocurrents in excess of 8 mA/cm² are required, using materials that require less than 0.8 V bias, exhibiting much greater than 10-year lifetimes, that can be deposited and incorporated into reactors which cost less than $100/m².

8.5.9 Material Parameters

Several determinants for the choice of materials will contribute to a commercially viable device. These include

- High photocurrent leading to high efficiency
- High photon capture
• High photon-to-electron conversion efficiency
• Low electron–hole loss
• $1.6 \text{ eV} < \text{band gap} < 2.2 \text{ eV}$ matching hydrogen and oxygen evolution potentials to minimize bias
• Low cost of source material
• Abundant supply
• Durability in aqueous environment
• Resistance to high solar irradiance
• Resistance to high temperature
• Electrode structures supporting high current flows
• Avoidance of poisonous or contaminating materials
• Electrolyte choice, ideally, commonly available, nontoxic, noncorrosive

8.6 Other Issues

Only issues of relatively direct impact on PEC research are touched on here. These include the following.

8.6.1 Hydrogen Storage

Sometimes seen as an “Achilles heel,” the problem of how to store and transport hydrogen remains to be solved. Historically, heavy steel cylinders have been and are still used to contain the gas at a range of pressures which ensure that the volume is not unacceptably large.

Only a small number of companies specialize in this field, and the development of much lighter weight containers based on carbon fibers, special resins, and so on is now well advanced, ready to supply the needs of the automotive industry, where the overall weight of gas plus container is a constraint.

Carbon nanotubes are being investigated, and metal hydrides are already available. Other solutions based on storing hydrogen in another form are also under review. Ammonia has its advocates, being a material which is handled in bulk by the fertilizer industry, and is composed of just two elements – N and H. However, transforming the hydrogen into and then back out of this chemical is troublesome, and other solutions are sought.

One interesting alternative is formic acid ($\text{H}_2\text{CO}_2$), as the $\text{CO}_2$ component may be usefully abstracted from power plants.
8.6.2 System Costs

To achieve the stated aim of competing with existing energy sources, without reliance on subsidy or tax hedge, the relevant benchmark is that of hydrogen production by SMR. This is a moving target, as the source for SMR is natural gas, itself a hydrocarbon under pressure from rising demand, and dwindling supply. However, we can compute the following using some assumptions:

- A 10% efficient device with bias requirements less than 0.8 V
- A 15-year minimum life
- A capital cost for the system (PEC + PV) of $160/m²

With these assumptions PEC hydrogen could cost less than $3/kg. A clear statement of aims such as these is essential for the development of a commercially convincing product.

8.6.3 Solar Concentration and Tracking

Modest solar concentration, in the $3 \times$ to $10 \times$ range, may be helpful in boosting output with a low capital cost increase, provided a nonsolar tracking configuration is adopted.

High concentration schemes are needed for high-temperature applications, but in the context of PEC cannot be recommended. The extra cost of the tracking component is likely to put the overall energy cost beyond an acceptable limit. Remember too that solar tracking implies clear beam radiation, thus reducing the number of locations where such applications can be sited.

8.6.4 The Issue of Intellectual Property (“IP”)

The world of “risk capital” funds, hedge funds, and the like places great reliance on the protection of IP by patents, and typically a family of patents is sought, covering all aspects of the given technology. Without seeking to detract from this viewpoint, it may be more important to rely on “know-how” and practical experience. For example, the original “dye-cell” patent has already expired, but the dye-cell team is busier than ever with applications for their know-how. However, there is a conflict between the wish for researchers to publish their work, and the need to maintain secrecy until patent cover has been applied for.
8.6.5 Further Issues

Maintenance/support in operation – the simplicity or otherwise of the PEC solution will affect the ongoing costs – what elements will need cleaning, and at what intervals? What elements will need replacement and at what frequency?

Hydrogen transportation from point of production to point of use will be an issue, especially since every pipe joint represents a potential leak of this finest of all gases.

8.7 Route to Market

Section 8.2 showed the huge sector represented by transport energy. For transport fuel, large scale and large company development is needed to make an impact. This will only happen when the requirements of cost, convenience, and safety have been met. At this point all the other uses of conventional fuels will be opened too.

That said, fleet vehicles tied to a refueling base (taxis, buses, delivery vehicles) avoid some of the “chicken-and-egg” problems associated with the eventual need for fully distributed infrastructure, while small fleet vehicles (fork-lift trucks, airport tugs) are even more readily serviced.

8.7.1 Other Markets

There are as many potential markets as there are activities in this world. The broad divisions noted above can also be subdivided into “grid connected” and “remote” or “off-grid” locations. The latter is of interest in that the price of energy at remote sites is already much higher than where grid connection exists. Niche markets alone do not justify great expenditure, so there is a need to think big – once a competitive process has been developed.

8.7.2 A Comment on “growth”

The idea that we need our world economy to grow for it to be healthy must be treated with caution. If population grows, then growth of supplies is needed. Apart from that, it would seem that the economic and banking crisis of 2008–2009 shows us that growth is chiefly needed to satisfy greed, and leads to collapse. Sustainability is a much more relevant criterion.
8.8 Conclusions

The opportunity to find a way to store solar energy poses huge challenges, but offers a huge reward – not just financial, but deeply satisfying in terms of achievement in an entirely new field for humanity. Clear targets must be set, ideally leading to system costs that are easy to accept in being no more than the fossil fuel alternative. Here, proposed is a figure of under $3/kg of hydrogen produced. To achieve this, we see that

- More efficient materials are needed, and a target of 8 mA/cm\(^2\) is set (to reach >10% efficiency)
- Bias requirements must be reduced, with minimal resistive voltage drops
- Durability must be increased to 15, or better 20 years
- System costs must be reduced, and a price not exceeding $160/m\(^2\) for both PEC and PV is set

By splitting water directly – in one step –the need for increased independent electricity generation is avoided, as is the need for grid connection. More importantly, the way would be opened to a real reduction in CO\(_2\) emissions at the most basic level. The tasks involved are large, and higher levels of funding are likely to be needed, especially if results are to be achieved in the near term.

References

Chapter 9
Emerging Trends in Water Photoelectrolysis

Scott C. Warren

9.1 Introduction

Manipulating the structure of matter and the distribution of optical energy at the nanoscale will be crucial to future progress in water photoelectrolysis. Fortunately, the rapid pace of developments in nanomaterials and nanophotonics provides ample inspiration as scientists continue to confront the long-standing challenge to invent an efficient and inexpensive way to store solar energy in chemical bonds of a fuel such as hydrogen.

Preceding chapters have outlined the challenging and often competing criteria that semiconductors must meet to achieve efficient and economical water photoelectrolysis. For a device based on a single photoelectrode to be successful, these criteria must include the following:

- A bandgap between approximately 1.8 and 2.4 eV
- Aqueous stability
- Band edges that straddle the water redox potentials
- Small ohmic losses for charge carrier transport
- Minority carrier mobility and lifetime that permit efficient charge transport to the semiconductor/solution interface
- Water oxidation/reduction kinetics that are faster than surface-mediated recombination

A few semiconductors have come close to meet most of these criteria, but none has met all. It seems that strategies that go beyond compositional variation are warranted. This chapter examines strategies for nanostructuring semiconductors as well as applying plasmonic effects to enhance device efficiency.
9.2 Nanomaterials

In recent years, interest in the use of metal oxides as photoelectrodes has greatly increased as significant gains in efficiency have been demonstrated. The interest in metal oxides also arises from challenges facing the aqueous stability of nonoxide semiconductors. Some classes of metal oxides that have inspired recent interest include spinels and delafossites, both of which offer highly tunable electronic and optical properties. But complex metal oxides often have low mobility and short carrier lifetimes, which has so far limited performance in devices based on these materials. Thus, nanostructuring is being increasingly explored as a route to minimize recombination.

In this discussion of nanomaterials, conditions are identified for building efficient devices from semiconductors with poor mobility and short carrier lifetimes. This approach is then analyzed within the context of anodized nanopore and nanotube arrays and directions are suggested for further improvement for this class of photoelectrodes.

9.2.1 Access to Small Length Scales

As new metal oxide semiconductors are identified and explored for water photoelectrolysis, a persistent problem that arises is the low drift mobility of electrons and holes and the short lifetimes of photogenerated carriers. Many materials that seem promising on the basis of bandgap and stability are found inadequate for photoelectrolysis because the low mobility of minority carriers prevents rapid charge transport to the aqueous interface. Metal oxides are also prone to nonstoichiometry, which can result in traps that promote recombination and shorten carrier lifetimes. When short lifetimes and low mobility arise in the same material, device efficiency can drop rapidly and photocurrents are far below those that might be otherwise expected.

Within the last decade, new nanostructuring approaches have emerged that suggest ways to solve these problems. This section reviews the major challenges that semiconductors face with low mobility and short carrier lifetimes and highlights the emerging strategies for addressing these shortcomings.

A central idea in the discussion of charge transport is that of the polaron. It is easiest to understand polarons – of which the two most common types are either large or small – by imagining an uncharged lattice with atoms located at their equilibrium positions. In this thought experiment, an electron or hole is added to the lattice and the lattice responds to this charge by delocalizing it over many atoms (a large polaron) or by localizing it over just a few atoms (a small polaron), as shown in Fig. 9.1. Because the additional charge fills or empties states with bonding or antibonding character, the charge influences bond lengths and angles. If the charge is delocalized the influence is small because it is spread over many atoms, while a highly localized charge
dramatically alters bond lengths and angles because it induces comparatively large changes in electron density. The concept of polarons, then, is used to indicate that the charge carrier and accompanying lattice distortion are coupled together and must therefore be considered as a single unit or quasi-particle.

Holstein has argued that the concept of self-trapping is probably the most important idea arising from the theory of polarons [1]. When a charge carrier is localized as a small polaron, the lattice shifts to new equilibrium positions within a period of a few lattice vibrations [2]. This increases the stability of the state, creating a deeper potential well that must be overcome to transfer the polaron to a neighboring atom. In the context of Marcus theory [3], it is said that there is a large reorganization energy associated with small polarons. It has also been argued that it is more appropriate to distinguish small and large polarons by self-trapping, rather than the extent of lattice deformation [4].

While the phenomena that distinguish small polarons from large polarons has been qualitatively outlined, there are also quantitative measures. Perhaps the most intuitive is electron or hole drift mobility. Because of the larger activation energy associated with transport, small polarons have lower drift mobility than large polarons. Bosman and van Daal have derived estimates of the limit for small polarons by assuming that self-trapping occurs when the persistence of a polaron on a site is much longer than the time period of a lattice vibration [2]. This allows an upper limit for small polaron mobility to be estimated as 0.1 cm²/V s. A lower limit for large polaron mobility was estimated as 0.8 cm²/V s, although the value was sensitive to the model’s assumptions. Other estimates have placed the lower limit for large polarons at 10 cm²/V s and the upper limit for small polarons at around 10⁻² cm²/V s [4]. Another distinguishing characteristic between small and large polarons is that the mobility of small polarons increases with temperature because transport is assisted by lattice vibrations; large polarons show decreasing mobility with temperature because lattice vibrations scatter carriers [4].

Fig. 9.1 Illustration of polaron size. The periodic potential, \( V(x) \), is plotted with the range of conduction band potentials (energies) depicted in gray. The wavefunctions, \( \psi \), of a small (solid line), intermediate (dashed line), and large (dotted line) polaron are plotted as a function of a single spatial dimension, \( x \). Illustration is based on [5].
To consider the impact of low mobility on device efficiency, it is first necessary to present the principal models of a Schottky barrier in a semiconductor–electrolyte junction. The most common approach is that of Gärtner [6], who calculated the photocurrent as the sum of all photogenerated carriers in the space charge layer and all photogenerated carriers produced in the bulk of the semiconductor that diffuse to the space charge layer. In this approach, emphasis is therefore placed on determining the minority carrier diffusion distance \( L_m \) (which is a function of mobility), the absorption coefficient \( \alpha \), and the space charge layer width \( w \) [7]. The Gärtner model assumes that no recombination occurs in the space charge layer, but this may not be suitable for materials with low mobility, where lowering the mobility will increase the time required for a minority carrier to cross the space charge layer and therefore increase the likelihood of recombination.

Seeing this limitation, Jarrett proposed a revised model that allows for recombination in the space charge layer [8]. Jarrett considered the effects of low mobility on device efficiency by first examining the influence of mobility on the transit time, \( \tau_t \), the time required for minority carriers to cross the space charge layer. The transit time is calculated as

\[
\tau_t = \frac{\varepsilon \varepsilon_0}{eN_i \mu},
\]

where \( \varepsilon \) is the static dielectric constant, \( \varepsilon_0 \) is the vacuum permittivity, \( e \) is the electric charge, \( N_i \) is the ionized impurity level concentration, and \( \mu \) is the mobility [8]. This is plotted in Fig. 9.2 for both small and large polaron transport for varying values of \( N_i \) with \( \varepsilon = 100 \), a typical value for a metal oxide semiconductors. For semiconductors with small polaron transport to have \( \tau_t \) below 10 ps, ionized impurity concentrations must be above \( 10^{19} \), which implies doping densities in excess of 0.1 at.%. An interesting feature is that \( \tau_t \) is independent of band bending, \( V_B \), because the higher carrier velocity that accompanies increased \( V_B \) is exactly countered by the greater space charge layer width.

There should be some caution in broadly applying (9.1) to all types of carrier transport at interfaces. For example, the relationship does not accurately model the transit time of ballistic transport because the calculation of \( \tau_t \) depends on the mobility, which is only accurate in so far as it measures a diffusive process, i.e., one that involves multiple scattering events [9]. Because the small polaron conductors have transport mediated by lattice vibrations, numerous scattering events will occur as the carriers cross the space charge layer. Therefore, the transit times as calculated by (9.1) should be representative of the behavior for this class of materials [10].

In calculations of efficiency, \( \tau_t \) is to be compared with the carrier lifetime in the space charge layer, \( \tau_R \). From a qualitative standpoint, if \( \tau_t > \tau_R \), a sizable proportion of minority carriers generated inside space charge layer will recombine before making it to the semiconductor/electrolyte interface. Taking a quantitative approach, Jarrett solved the diffusion equation for an illuminated planar semiconductor/electrolyte interface and calculated incident photon-to-current efficiency (IPCE) for systems with varying \( \tau_t/\tau_R \) [8]. One series of calculations used as
input the conditions that $zL = 1$ and $s/k = 1$, where $z$ is the absorption coefficient, $L$ is the root-mean-square distance traveled in time $t_\tau$ in the absence of a field, $s$ is the rate of transfer of minority carriers from the semiconductor surface to the electrolyte and $k$ is the rate of transfer of carriers from within the semiconductor to its surface. These conditions state that most light is absorbed in the space charge layer of the semiconductor (because $L$ must be less than $w$) and that every minority carrier that makes it to the semiconductor/electrolyte interface contributes to the photocurrent, i.e., there is no surface recombination. Fig. 9.3 shows the results of these calculations, which indicate that the IPCE is dramatically influenced by band bending and $t_\tau/t_R$. In every circumstance, low efficiencies are obtained for very small degrees of band bending because there is an insufficient field to support charge separation. But with modest band bending, high IPCEs (over 80%) can be obtained only when the transit time is similar to the recombination time.

From the foregoing arguments, it becomes clear that small polaron-type semiconductors with $t_R$ in the picosecond range need high doping levels ($>10^{19}$ cm$^{-3}$) to achieve high efficiencies. This implies a narrow space charge layer width, $w$, as calculated by
where $N$ is the carrier concentration. The relationship between $w$ and $N$ is included in Fig. 9.2 for a semiconductor with $\varepsilon = 100$ and $V_B = 0.25$ V, the latter is a typical value for an operating photoelectrochemical device. In Fig. 9.2, it is seen that small space charge layer widths are necessary for semiconductors with small polaron transport if small transit times are desired. For example, under the stated conditions with $N = 10^{19}$ cm$^{-3}$, the space charge layer width is 16 nm; for $N = 10^{20}$ cm$^{-3}$, the space charge layer width is 5 nm.

With such small space charge layer widths, it is likely that a significant proportion of light is not absorbed within the space charge layer. Additionally, within the Jarrett model, the absorption coefficient $\alpha$ is to be compared with $L$, such that high efficiencies are only achievable when $\alpha L$ is equal to or greater than one. But by increasing the doping level to decrease $\tau_t$, $L$ also decreases. For example, a small polaron semiconductor with a mobility of 0.1 cm$^2$/V s and a doping level of $10^{19}$ cm$^{-3}$ has a transit time of 55 ps. If this time is similar in magnitude to the lifetime in the space charge layer, then $\alpha L$ is the relevant parameter (not $\alpha w$) and it must be much less than one since $L$ is just 1.1 nm. Jarrett calculates that in going from $\alpha L = 1$ to $\alpha L = 0.1$, the photocurrent efficiency lowers by a factor of 2.5.

![Fig. 9.3](image-url)
Something interesting happens, however, when the doping level increases further. If, through increased doping, $\tau_t$ becomes significantly less than $\tau_R$, recombination in the depletion layer can be neglected and consequently the Jarrett model simplifies to the Gärtner model. This is illustrated in Fig. 9.4, where the IPCE is calculated for varying doping levels. It is seen that at low doping levels ($\tau_t < \tau_R$) the Jarrett model applies and predicts low IPCEs because of the significant recombination in the depletion layer. At intermediate doping levels ($\tau_t$ is near $\tau_R$) IPCEs reach their peak value because recombination in the space charge layer is diminished. Finally, for high doping levels ($\tau_t < \tau_R$), the Gärtner model applies because recombination in the space charge layer is negligible. But as doping increases, IPCEs decrease because an increasing proportion of photons are absorbed outside of the narrowing space charge layer. This would suggest that there is an optimum doping density where IPCE is maximized, as seen in Fig. 9.4.

The foregoing calculations by Jarrett and Gärtner are modeled on planar semiconductor–electrolyte interfaces and, as a consequence, low IPCEs are predicted for semiconductors with low mobility and short lifetimes. But if nonplanar geometries are explored – specifically, nanostructures in which the semiconductor fits entirely within the space charge layer – it should be possible to make efficient devices. Specifically, the successful approach will involve doping the semiconductor to the point that $\tau_t < \tau_R$ and the Gärtner assumption of no space charge layer recombination applies. If the material can be designed such that all photons are absorbed in the space charge layer – a difficult geometric constraint, considering the small widths as seen in Fig. 9.2 – IPCEs approaching 100% are feasible. Of course, this also assumes that the semiconductor is optically thick such that all above-bandgap photons are absorbed.

This discussion has assumed that $\tau_R$ is independent of the doping concentration and the position of the carrier within the depletion layer. In fact, both of these parameters influence $\tau_R$, because the recombination rate depends on the
concentration of charge carriers in the bands and the relative occupation of electrons and holes in traps [11]. In general, band-to-band recombination is dominant in direct bandgap semiconductors while trap-mediated recombination is dominant in indirect bandgap semiconductors [9]. Minority carrier lifetime for band-to-band recombination is inversely proportional to the majority carrier concentration $N$. In this case, increasing the doping density does not significantly alter the ratio of $t_t$ to $t_R$ since both have the same functional relationship to $N$. On the other hand, band bending transports excess electrons and holes away from each other, which minimizes direct band-to-band recombination within the space charge layer.

The situation is more complicated for trap-mediated recombination. When traps are located near the middle of the bandgap and light intensities are low (the situation in which there are fewer photogenerated carriers than majority carriers), minority carrier lifetimes are independent of $N$ and inversely proportional to the trap concentration [9]. The case for traps closer to a band edge is more complex, as shown in Fig. 9.5. Shockley and Read discussed four regimes of recombination behavior under steady-state conditions for this type of system [11]. First, for highly doped p-type semiconductors, the rate-determining step is electron trapping because the very high concentration of holes leads to a short $t_p$. In this regime, $t_e$ and therefore $\tau$ is relatively insensitive to changes in doping. Second, for semiconductors with modest levels of p-type dopants, the rate-determining step is now the trapping of holes because of their lowered concentration. Here, $t_p$ (and $\tau$) change with doping as $1/N$. Third, for modestly doped n-type samples, traps are still
mostly unoccupied and the even lower concentration of holes makes their removal the rate-limiting step. But in this regime, increasing doping fills the traps which decreases the hole lifetime; as a consequence $\tau$ changes with doping as $1/N$. Fourth, with all the traps filled by electrons, the rate-determining step is hole trapping. Lifetime is now independent of doping because increasing the doping level further does not increase the concentration of electrons in traps. As shown in Fig. 9.5, only in the range of modest doping does lifetime scale as $1/N$.

These four regimes were derived based on the assumption of low light intensities. As long as the majority carrier concentration does not change significantly upon light irradiation, the light intensity is considered low. For the regimes of high doping, this condition should hold for typical light intensities. It should also be noted that lifetime was calculated assuming the absence of a field. In the presence of a field, however, lifetimes should be further increased because the field separates photogenerated carriers. With the effect of the field accounted for, $\tau$ can be equated to $\tau_R$.

Two additional effects can influence device efficiency – in these cases, by altering $\tau_t$. In general, carrier mobility decreases with increasing impurity concentration because of impurity scattering [9], although the relative importance of impurity scattering in materials of low mobility – which already have mean free paths on the length scale of atomic distances – is probably minimal. It is also found that mobility changes under strong electric fields, which is a condition that arises within the space charge layer at high doping levels. For a potential drop of 0.25 V across a 25-nm space charge layer, the average electric field is 100 kV/cm. Semiconductors with high mobility typically have reduced mobility under high fields, although the situation is not straightforward for materials with low mobility. For example, if the average time required for a carrier to hop between centers becomes comparable to the lattice relaxation time, mobility can increase because self-trapping is diminished [12]. In fact, in some materials carriers can shift between small and large polaron transport [13]. It is therefore difficult to make any general predictions and the behavior of mobility under high fields must be determined experimentally for each case.

To summarize the foregoing arguments, semiconductors with small polaron transport and short carrier lifetimes can be used to make efficient devices only if several design criteria are met. First, doping levels must be sufficiently high such that $\tau_t < \tau_R$. This will eliminate most recombination within the depletion layer. Second, light absorption should be confined to the space charge layer. Based on the arguments presented, this condition is unlikely to be met for electrodes with planar geometries. This can be achieved, however, if the semiconductor is suitably nanostructured. By designing the material such that the electrolyte interface is never more than 5–10 nm from any place where minority carriers are generated in the semiconductor, high IPCEs are possible.

In the discussion that follows, a synthetic method that exploits nanostructuring concepts is presented, allowing these design criteria to be met. Several compelling synthetic strategies have emerged in recent years and further progress may require the application of these strategies to new metal oxide materials. The discussion
addresses anodization of metals to produce nanotube arrays. It has already been seen that the application of these synthetic techniques to photoelectrochemical devices leads to improved performance.

### 9.2.2 Nanopore and Nanotube Arrays by anodization

It has long been known that the anodization of certain metals leads to the production of porous metal oxides and hydroxides. These metals include Al [14–16], Ti [17, 18], Ta [19], Cd [20], Nb [21], Mg alloys [22], W [23], Sn [24], Fe [12, 25], Ag [26], and Si [27]. Only recently conditions have been identified that allow the formation of well-controlled, uniform structures. For example, in 1995 an aluminum anodization process to develop hexagonally packed pores with ordered domains was developed [16]. In 2001, it was discovered that anodization of titanium foils led to the production of TiO$_2$ nanotube arrays [18]. More recently, a anodization process to form nanoporous [25] and nanotubular [28] iron oxides was developed. Examples of these materials are shown in Fig. 9.6. Pore sizes in the range of 20–100 nm are
typical for this class of materials, although some materials, such as aluminum, have achieved even greater variations [31]. In the case of the nanotubular materials such as TiO$_2$ and Fe$_2$O$_3$, wall thicknesses in the order of 5–30 nm are typical.

In the context of water splitting, TiO$_2$ nanotube arrays have attracted the greatest attention. TiO$_2$ is an excellent candidate for nanostructuring because of its low room-temperature mobility and modest carrier lifetime. TiO$_2$ is generally n-type and therefore the hole mobility is the relevant parameter in determining $\tau_t$. Only a few experimental [32–34] and theoretical [35, 36] determinations of hole mobility have been performed. Recent calculations of room temperature hole mobility in rutile provide values of $5 \times 10^{-3}$ to $4 \times 10^{-5}$ cm$^2$/V s for the most favorable (adiabatic) modes of charge transfer, while in anatase only one adiabatic mode was found, yielding a mobility of $2 \times 10^{-3}$ cm$^2$/V s [36]. Where these calculations could be compared with experiment – high temperature conductivity measurements on rutile [33] – good agreement was found. Models of lattice distortion arising from the hole polarons were also calculated, as shown in Fig. 9.7. These showed that although the distortion influenced bond lengths up to several atoms away from the hole, these long range distortions were small in magnitude and had little influence on the hole structure and hopping energies. The large reorganization energies in both anatase and rutile (mostly between 2.0 and 2.5 eV) suggested that the hole transport be classified as a small polaron [4].

Hole lifetime in TiO$_2$, as in all materials, is sensitive to impurity concentration and identity, grain boundaries, defects, surface states, electric fields, and the concentration of photogenerated carriers. For this reason, determination of hole lifetime under one set of conditions may vary substantially from another set. Studies of TiO$_2$ nanoparticles (of diameter ~2 nm) showed electron–hole lifetimes ($\tau$) of 30 ns [37, 38]. With the cautious approximation that $\tau_R$ of TiO$_2$ nanotube arrays is also 30 ns – probably an underestimate of $\tau_R - (9.1)$ shows that a doping level of approximately $1 \times 10^{18}$ cm$^{-3}$ is necessary to have $\tau_t$ equal to $\tau_R$. This corresponds to a space charge layer width of 55 nm. As suggested by the trends discussed above, decreasing $\tau_t/\tau_R$ further diminishes

![Fig. 9.7 Distortions in bond lengths, in Ångströms, surrounding a hole “H” on an oxygen atom in (a) rutile and (b) anatase. The images show just part of the cluster that was used for the calculation. Reproduced with permission from [36]. Copyright 2009 American Chemical Society](image-url)
recombination in the space charge layer. This speaks, therefore, to the opportunities in nanostructuring TiO$_2$. In practice, because the lifetimes in TiO$_2$ are relatively long, much of the work on TiO$_2$ does not seek to explicitly dope the material for the sake of improving charge separation. Rather, the material is doped only modestly as an n-type semiconductor through the formation of oxygen vacancies and the incorporation of contaminants from either the titanium foil (the typical impurity content of titanium foil used to make the nanotubes is 0.1–0.3%) or the electrolyte.

In 2004, Grimes reported the application of TiO$_2$ nanotube arrays to water photoelectrolysis [39] and in 2005 provided an improved synthetic method that produced 224-nm long anatase nanotubes with a wall thickness of 34 nm. This led to an IPCE of over 90% at an applied bias of 1 V when irradiated at 337 nm [40]. In the later work, it was suggested that a key aspect of the improved performance was that the space charge layer would fit entirely within the nanotube walls, thereby implying that electric field-mediated charge separation was important to achieve high efficiencies in these nanostructures. The best performing materials synthesized using TiO$_2$ nanotubes converted light energy to hydrogen energy at an efficiency of 16.25%, as shown in Fig. 9.8, although it should be emphasized that ultraviolet light and not the full solar spectrum was used [41].

Considering the large bandgap of TiO$_2$, materials exposed to sunlight produce relatively low photocurrents. When AM 1.5 is used to photoelectrolyze water on titanium dioxide nanotubes, the maximum current densities are 1.1 mA/cm$^2$ with conversion efficiencies of 0.3% [42]. Numerous efforts have been made to dope TiO$_2$ to increase its response to visible light [43–46]. While visible light absorption can be induced in these materials and this visible light reaction can be used to drive photochemical reactions at the nanoparticle surface, efficiencies from visible light absorption are usually low and efficiencies from ultraviolet light typically decrease as well. The decreased efficiencies often arise from the creation of localized states within the bandgap that act to trap electrons or holes.

Fig. 9.8 Current densities and photoconversion efficiencies for TiO$_2$ nanotube calcined at either 525 or 580°C. Reproduced from [41] with permission of IOP Publishing Ltd
An approach that has been successfully implemented to improve light absorption in TiO₂-type nanotubes has been the incorporation of either iron or copper. A thin metallic film is sputtered that contains a mixture of titanium and iron, which is then anodized. This leads to the growth of nanotubular metal oxides with compositions that depend on the starting Ti:Fe or Ti:Cu ratio. In the case of iron, incorporating 6% into titanium provided the best results, where visible light photocurrents were obtained. Compared to the Ti-only materials, however, the sample had to be biased to a much greater degree to see similar photocurrents: onset potentials were nearly 1 V more positive. Nevertheless, a plateau photocurrent of 2.0 mA/cm² was achieved under AM 1.5, nearly twice that of the TiO₂-only materials. The large bias most probably resulted from the fact that the iron provides trap sites well below the TiO₂ conduction band edge, implying that large biases would be needed to free these trapped electrons.

A similar approach was used to incorporate copper into TiO₂-based films. Metallic Cu–Ti films with Cu contents of 60–74 at.% were anodized and subsequently calcined. This led to the production of nanotubular arrays composed of CuO, anatase, and rutile. The films exhibited p-type behavior, arising from CuO. This allowed a two-photoelectrode device to be built comprising a p-type CuO–TiO₂ nanotubular array and an n-type TiO₂ nanotubular array, as illustrated in Fig. 9.9. As the device employed two photoelectrodes, sufficient bias was established to drive unassisted water photoelectrolysis (Fig. 9.9b). Photocurrents of approximately 0.25 mA/cm² were established.

More efficient devices will require the improved utilization of visible light at both photoelectrodes. As suggested in the foregoing discussion, doping TiO₂ provides a compelling route to achieve this. Yet, it is well established that the doping also leads to an increased degree of recombination depending on conditions (dopant concentration, identity, redox behavior). The seminal studies of Hoffmann indicated, however, that the use of certain dopants in extremely small particles can, additionally, prolong lifetimes and facilitate interfacial charge transfer [43]. This effect was lost when particle sizes were increased. This therefore suggests that doping TiO₂ strategies that can decrease τ_e can also be used to enhance light absorption and
charge transfer if implemented in materials with sufficiently small particle sizes. These sizes are increasingly being approached with synthetic methods such as anodization, which may permit TiO$_2$ to both absorb increasing amounts of visible light while minimizing the likelihood that holes recombine with electrons.

Iron oxide ($\alpha$-Fe$_2$O$_3$) is another well-known material that can benefit from a combined doping and nanostructuring approach, as outlined above. Hematite has been variously classified as either a small polaron conductor or a borderline case between small and large polaron behavior [2]. Some of the discrepancy probably arises from its highly anisotropic hole transport, with one experimental study on heavily doped p-type hematite showing that mobility within (0001) being somewhere between a factor of 10 and 10,000 higher than in [0001] [48]. A theoretical study has predicted a 1,000-fold difference in conductivity [49]. Other studies have shown that the anisotropy is highly dependent on dopant identity and concentration [50]. The most careful measurements of hole and electron transport in poly-crystalline hematite have indicated a mobility of 0.2 cm$^2$/V s at 1,400 K and 0.1 cm$^2$/V s at 1,200 K, respectively [2]. Theoretical studies, in particular, have suggested that the transport mechanism is based on small polaron in spite of bond distances being distorted over significant distances [51].

Lifetimes in hematite are much shorter than TiO$_2$. Carrier relaxation to the conduction or valence band edge occurs within a few hundred fs, and recombination or trapping occurs within 3–5 ps. To have $\tau_t = \tau_R$, the doping density should be $10^{20}$, or approximately 0.25 at.%, assuming that these values for lifetime are similar to $\tau_R$. This corresponds to a space charge width of 5 nm for a potential drop of 0.25 V. Therefore, it would be desirable to synthesize highly doped hematite nanostructures with dimensions of less than 10 nm.

Investigations of iron anodization have indicated that it is possible to synthesize porous, amorphous iron oxide with a nanotubular morphology (Fig. 9.6) [25, 28]. Heat treatment crystallizes the iron oxide to form hematite. Recently, anodization conditions were found which lead to the formation of nanotubular hematite with wall thickness of approximately 6 nm [28]. The best performing nanotubes were formed under a reducing environment, which should create a high concentration of oxygen vacancies. It has been shown that oxygen vacancies create a donor level about 0.20 eV below the conduction band [2], thereby serving as an electron trap when empty. The resulting material exhibits photocurrents of 1.41 mA/cm$^2$ at 1.5 V vs. RHE, although it should be noted that the sample was only illuminated with 87 mW/cm$^2$ at AM 1.5. There are better dopants for hematite than oxygen vacancies since the vacancies also act as traps, and achieving a controlled doping density would further lead to improved control over $\tau_t$. Nevertheless, the initial studies with $\alpha$-Fe$_2$O$_3$ represents promising initial investigations of a compelling and versatile nanostructuring technique. An interesting example of this combined doping and nanostructuring approach has been demonstrated for Si-doped hematite prepared by chemical vapor deposition [52]; if applied to the nanotube structures significant performance gains may be possible.

To apply this as a fabrication technique for new metal oxide compositions, some understanding of the formation mechanism must be had. There are numerous
studies that outline the mechanism of the formation of such anodized films and these studies have revealed that there are many similarities in spite of differing compositions [17, 53]. The anodization process occurs as follows:

1. Prior to anodization, there is an oxide layer of varying thickness on the metal surface. Application of an anodizing potential causes electric fields to be concentrated in thinner regions of the oxide layer.
2. The anodizing potential causes the reaction $\text{M}^0 \rightarrow \text{M}^{n+} + ne^-$ to occur at the metal/oxide interface, with the reaction occurring at the greatest rate where electric fields are most concentrated. The high-electric field drives the metal ion to the metal oxide/electrolyte interface.
3. The anodizing potential drives water decomposition ($\text{H}_2\text{O} \rightarrow 2\text{H}^+ + \text{O}_2^-$) and, possibly, electrolyte decomposition at the electrode/solution interface. The oxygen anions diffuse through the oxide layer to form the metal oxide. Some proportion of anions from the electrolyte are also incorporated into the oxidized material. This metal oxide layer through which metal and oxygen ions diffuse is called the barrier layer.
4. The water decomposition reaction increases the concentration of electrolyte species that etch the metal oxide – usually this is $\text{H}^+$ or $\text{F}^-$. Because these species are most concentrated in the vicinity of the concentrated electric field, etching occurs most rapidly in these locations, which leads to pore formation.
5. A self-reinforcing process develops in which the electric field becomes more concentrated in the pore bottoms because this is where the barrier layer is the thinnest. This promotes steps 2, 3, and 4, which allows the anodization and pore formation to continue.
6. The formation of an ordered pore structure arises only when a careful balance of oxide growth and dissolution are achieved [54, 55]. As the metal is oxidized, it also expands, which induces stresses. Modest stresses force the growing pores to organize in a hexagonal array so as to provide the most efficient packing of the expanding oxide [54, 56, 57]. Temperature, electrolyte, and voltage are the parameters typically used to control growth and dissolution rates.

Figure 9.10 outlines these steps for the anodization of nanoporous and nanotubular morphologies. The nanotubular structures can be seen as containing two sets of pores: one large pore inside each tube and a second pore that surrounds the outside of the tube. In contrast, the nanoporous structures only develop the single large pore. It has been proposed [17] that variations in ion mobility in different metal oxides determine whether nanotubular or nanoporous morphologies form. For example, Ti and O ions have low mobility, which implies that the oxide layer at the bottom of the pores will be relatively thin under steady-state conditions. Because of the relatively thin oxide layer, some Ti will remain unoxidized within the pore walls. The electric field then becomes sufficiently concentrated in these unoxidized regions for electric field-assisted dissolution to occur. This is a self-reinforcing process that continues throughout the anodization and leads to nanotube formation.

An important point is taken from the work on Cu–Ti systems [47]. While copper has not yet been found to be suitable for anodization as a pure metal, it has been
successfully implemented in the form of Cu–Ti alloys with Cu compositions up to 74 at.%. This would therefore suggest that mesoporous complex oxides can be fabricated by anodization as long as one component of the starting alloy can be anodized. Because the resulting oxide is most often amorphous [58], heat treatment that leads to a crystalline, homogeneous complex oxide is critical. Based on the extensive experience in heat treatments of TiO$_2$ nanotubes, annealing temperatures should take into consideration the possibility of structural collapse.

The controlled incorporation of dopants is another important direction for further improvement in anodized materials. Suitable methods may include direct incorporation of a metal ion dopant during preparation of the metal film or by incorporation of anions from the electrolyte during anodization.

With such material properties suitably controlled, it may be possible to fabricate efficient devices made from complex oxides with poor mobility and short carrier lifetimes.

### 9.3 Nanophotonics

In addition to improved methods for controlling the structure and composition of semiconductors, there is emerging work in the field of nanophotonics that may be favorably exploited for use in water photoelectrolysis. Much of the work in nanophotonics seeks to manipulate photons using the properties of nanostructured materials in ways that cannot be achieved with macroscopic optical components.
There are several emerging areas of nanophotonics, which include plasmonics, multiple exciton generation, photonic crystals, and upconversion. While these concepts have been applied to the design of solar cells, the exploration of these concepts in the context of water photoelectrolysis devices is limited. It is therefore of interest to consider the ways in which advances in nanophotonics could influence the direction of future work in water photoelectrolysis. This section emphasizes developments in plasmonics that may be useful if successfully applied to water photoelectrolysis.

### 9.3.1 Plasmonics

Much interest has been stimulated in plasmonics, which arises from the useful property of certain nanostructured materials to confine and manipulate information encoded within a signal of photons. This opens the prospect for building microprocessors capable of computing with optical signals rather than electrical ones, which could lead to breakthroughs in speed and efficiency of computing devices [59]. Other demonstrations within plasmonics include the ability to create lenses with resolution well below the diffraction limit of light, thus paving the way to optical microscopes with unparalleled resolution [60–62]. The underlying physics that are exploited in plasmonics are also of interest in solar energy conversion devices, since the ability to confine and scatter photons should improve photon absorption and utilization in water photoelectrolysis devices.

Already, several plasmonic effects have been explored within the contexts of thin-film silicon, dye-sensitized, and organic solar cells. The phenomena that have been observed and exploited – principally near-field absorption and light scattering – are described below. Together, these findings suggest strategies for further efficiency enhancements in water photoelectrolysis.

When a metal nanoparticle is subjected to an oscillating electric field such as the field intrinsic to light, the free electrons in the nanoparticle will respond to the field by oscillating. The response to the field is quantified by the polarizability of the metal. The polarizability $\alpha$ is given by

$$\alpha = 4\pi a^3 \frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m},$$

where $a$ is the particle radius, $\epsilon$ is the wavelength-dependent dielectric function of the metal, and $\epsilon_m$ is the dielectric constant of the surrounding medium. The equation indicates that the polarizability is greatest when the real part of $\epsilon$ is equal or close to $-2\epsilon_m$. This provides the condition for a strong resonance of the free electrons, which is called the dipole surface plasmon. Although most metals have resonances in the ultraviolet, gold and silver are most polarizable in the visible or infrared.
The interaction of light with the metal nanoparticle is defined by the extinction cross-section $C_{\text{ext}}$, which is the sum of the absorption cross-section $C_{\text{abs}}$ and the scattering cross-section $C_{\text{sca}}$:

\[
C_{\text{abs}} = k \text{Im}[\chi] = 4\pi k a^3 \text{Im} \left( \frac{\varepsilon - \varepsilon_m}{\varepsilon + 2\varepsilon_m} \right),
\]

(9.4)

\[
C_{\text{sca}} = \frac{k^4}{6\pi} |\chi|^2 = \frac{8\pi}{3} k^4 a^6 \left| \frac{\varepsilon - \varepsilon_m}{\varepsilon + 2\varepsilon_m} \right|^2.
\]

(9.5)

For small particle sizes ($<25$ nm diameter for gold), $C_{\text{abs}}$ is the dominant term and $C_{\text{sca}}$ can be neglected [63]. But with increasing particle size, $C_{\text{sca}}$ becomes larger, owing to the $a^6$ dependence on radius. Therefore, the behavior of the plasmon can be tuned with particle size.

When $C_{\text{abs}}$ is the dominant term, the plasmon – i.e., the collective oscillation of electrons – persists on the metal nanoparticle for between 5 and 10 fs before decaying as heat [64]. The nonrandom displacement of electrons produces an electric field that extends beyond the nanoparticle’s surface. When this electric field extends into a region of space that contains a semiconductor, an electron–hole pair can be generated in the semiconductor if the energy of the plasmon is greater than the semiconductor’s bandgap. The electric field decays with distance $d$ from the metal particle as $d^{-3}$ so that the probability of absorption in the semiconductor decreases with increasing distance from the metal nanoparticle. Thus, the plasmon effectively “focuses” photons to increase photon absorption in regions of the semiconductor that are close to a metal nanoparticle; this is termed near-field absorption.

Plasmonic near-field absorption has been modeled and exploited in silicon, organic, and dye-sensitized solar cells. For example, an organic solar cell achieved a $2.7 \times$ increase in photocurrent after incorporating copper nanoparticles [65]. A tandem organic cell exploiting the plasmon resonance of silver nanoparticles achieved a photocurrent of 2.5 mA/cm$^2$, while the identical device without the plasmon resonance was expected to achieve 1.9 mA/cm$^2$ [66]. Two examples of the modeling of the electric field induced by the surface plasmon are shown in Fig. 9.11. At the resonance wavelength, significant field enhancements are observed and the effect extends over 50 nm from the silver, albeit with decreasing field.

When $C_{\text{sca}}$ is the dominant term, the oscillating dipole primarily causes scattering of the incident light. In general, scattering with metal nanoparticles has been exploited to increase the optical path of photons through the absorbing layer or to couple light into the waveguiding modes of thin films. This was elegantly demonstrated in a thin film silicon solar cell in which 12 or 16-nm silver nanoparticles scattered light into waveguiding modes, leading to a 16-fold enhancement in absorption at certain wavelengths in the infrared [67]. Using a similar thin-film silicon geometry, the waveguiding modes were modeled, as shown in Fig. 9.12a. A strong enhancement of the field is observed, which is localized in
Fig. 9.11 Plots of field intensity for two types of solar cells. (a) A cross-sectional view of a 50-nm thick silicon solar cell separated by a 10 nm silica layer from a 100-nm wide silver nanowire. As compared to the incident field, a nearly tenfold increase is observed in regions of the silicon close to the silver. Reproduced from [68], with permission. (b) A cross-sectional view of 5-nm diameter silver nanowires sitting on a quartz substrate and in contact with copper phthalocyanine (CuPc). Nanoparticles are spaced 5-nm apart (d), which induces coupling between the dipoles and creates an increased field between adjacent nanoparticles. Contour labels represent the absorption enhancement factor as compared to the incident field. Reprinted with permission from [66]. Copyright 2004, American Institute of Physics.

Fig. 9.12 Models of scattering from metal nanostructures. (a) Same geometry as Fig. 9.11a showing a waveguiding mode. Reproduced from [68], with permission. (b) A model of a Si/Ag device with a 100-nm wide and 50-nm deep groove and (c) a simulation of the field scattered from the groove. Reproduced with permission from [69]. Copyright 2008 American Chemical Society.
the silicon layer [68]. An alternate scattering geometry is shown in Fig. 9.12b, with a thin-film silicon cell deposited on a silver substrate with a groove in the silver substrate. This geometry leads to a high degree of lateral scattering, mostly being confined to a 200-nm layer from the silver surface, as shown in Fig. 9.12c. In comparing the absorbance of silicon in the scattering geometry to a typical planar geometry, a broadband enhancement in absorbance was observed. At 700 nm the scattering film improved light absorption by a factor of 2, at 900 nm a factor of 1.2 was achieved, and at 1,100 nm it was improved by a factor of 2.4.

The efforts to enhance efficiency in solar cells through use of metal nanoparticles should provide a general guide for the application of plasmonic concepts to water photoelectrolysis, although differences in device design, such as the presence of an electrolyte, suggest that there may be additional criteria. When metal nanoparticles are located at the semiconductor/electrolyte interface, considerations of stability, Fermi level, and band bending are essential. Secondary considerations may include the influence of catalytic effects and the possible formation of trap states at the metal/semiconductor interface.

Near-field absorption enhancement can be used to confine light absorption to the proximity of metal nanoparticles. This could prove useful in a number of circumstances, including those described in Sect. 9.2, where the transport of minority carriers across the space charge layer competes with recombination. In such a case, the placement of nanoparticles at the semiconductor/electrolyte interface would lead to carrier generation at or near the interface, thereby limiting recombination losses. Alternatively, particle placement at the back interface could assist semiconductors with poor transport of majority carriers. In the case of scattering from metal structures, either of the designs that are illustrated in Fig. 9.12 may have promise, although variations in the dielectric function of the semiconductor can dramatically influence the efficacy of such structures [69]. In general, however, the possibility of confining light absorption to a decreased volume of semiconductor, as such designs allow, could allow a larger photovoltage to develop while potentially decreasing material costs.

### 9.4 Outlook

Future directions in water photoelectrolysis are myriad and there are many opportunities to discover cost-effective methods for the conversion of solar energy to chemical energy. Although an emerging trend in water photoelectrolysis is the development of nanostructured photoelectrodes, the full potential of nanostructures can only be achieved when the limitations of a specific material are accounted for, such as high recombination rates or poor charge transport. In particular, the directions suggested here provide a rationale for the development of highly doped nanostructures in which the feature size of the structure is commensurate with twice the space charge width. This strategy is expected to lead to significant improvements in efficiency for semiconductors with small polaron transport and
short lifetimes. With this design approach, many semiconductors – among the most promising are complex oxides – can realistically be considered as candidates for water photoelectrolysis.

The application of plasmonic metal nanostructures to water photoelectrolysis is an intriguing and potentially rich area for future exploration. By providing several general mechanisms to alter the traditional $e^{-x}$ absorption profile in semiconductors, the construction of new design rules for water photoelectrolysis devices is possible.

Acknowledgments S. C. W. thanks Michael Grätzel and Hen Dotan for fruitful discussions and the Swiss Federal Office of Energy (PEChouse, project number 102326) and the European Commission’s Framework Project 7 (NanoPEC, Project 227179) for support.

References


Index

A
Absorbance, 54
Absorbed photon-to-current conversion efficiency (APCE), 54
Absorption coefficient, 18
Accumulation layer, 35
Accumulation of holes, 98
Aliovalent dopants, 23
AM 1.5 global solar irradiance spectrum, 224
Anion dopants, 187
Anodization, 135, 302–308
Atmospheric pressure metal organic chemical vapor deposition (APCVD), 140–142
Automated electrochemical deposition system, 194

B
Band bending in a nanoparticle, 60
Band edge positions as a function of pH, 41 for selected semiconductors, 42
Bandgap, 18
Band structure, 167 of rutile TiO2, 16 of silicon, 15
Beer–Lambert law, 237
Benchmark for PEC water splitting, 9
Bias voltage, 284
BiFeO3, 162
Binary iron oxides, 169
BiVO4, 28, 162–168 crystalline phases, 163 density-of-states, 166 IPCE, 164 I–V curves, 168 morphology, 163
Brouwer diagram, 28
Butler–Volmer equation, 47

C
Calibrated photodiodes, 93–94
Capital cost of hydrogen, 284
Carrier lifetime, 296, 300
Cell constant, 77
Charge carrier mobilities, 19
Charge transfer, 57
Charge transfer current, 46
Chemical-bath deposition (CBD), 163
Chemical fuels, 6
Chemical solar-to-hydrogen (STH) conversion efficiency, 223
Climate change, 280
CO2 capture, 5, 6 emissions, 4 levels, 4 source of carbon, 6
Coal, 3
Co–Al–Fe oxides, 185
Cobalt-phosphate (Co-Pi), 148
Co2+ catalyst, 148
Co-doping, 29
Colloidal hematite, 132
Colloidal solutions, 177
Colored glass filters, 93
Combinatorial dopant optimization, 187–193
Combinatorial library, 181
Combinatorial methods, 180
Combinatorial screening, 180
Commercial challenges, 280
Commercially attractive targets, 287
Complex impedance, 102
Conductivity, 20
Conservation law’s for defect chemistry, 22
Constant phase element (CPE), 113
Conventional slurry (CS) method, 163
Cost per kg of hydrogen, 9
Cost sensitivity analyses, 218
Counter electrode, 75
Coupled photovoltaic–electrolysis systems, 9
Covalent bonding, 15
CuBi₂O₄, 169
Current and voltage relationships, 241–245
Current matching, 175

D
Dark current, 47
Decomposition potentials, 52
Deep depletion layer, 35
Deep donors and acceptors, 20
Defect-chemical reactions, 22
Defect equilibria, 27–28
Degeneracy factors, 20
Density-of-states (DOS), 19
Device configurations, 60–62
Diffusivity, free carriers, 57
Direct and indirect transitions, 17
Dissolution reactions, 23
DOE target, 279
Donor density, 34
Dopant energy levels, 25–26
Doped TiO₂, 26
Doping, 187
Doping of complex metal oxides, 28–29
DOS. See Density-of-states (DOS)
Dual-bed photocatalyst system, 217
Durability of PEC devices, 286
Dye-sensitized TiO₂, 9
Electroneutrality equation, 27
Electronic compensation, 23
Electronic structure, 15
Electron transfer theory, 45
Emerging trends in water photoelectrolysis, 293–313
End use of hydrogen, 282
Energy conversion efficiency, 53
demand, 278
diagram, 14, 39
infrastructure, 7
levels of dopants, 21
storage, 7
Equilibrium constant, 23
Equivalent circuit diagrams, 102
External bias potential, 40
Extremely thin absorber, 143–145

F
False color photocurrent image, 183, 184
Fe–Cs–Nd–Cu system, 184
Fermi level pinning, 44, 147
Fe–Ti–Si–Al system, 192
Fill factor, 242
First Law of Doping, 24
Fischer–Tropsch technology, 8
Flatband potential, 43
Fluorescent indicators, 181
Fossil fuels, 3
reserves, 5
Frank–Condon principle, 46
Free electron concentration, 19
Frequency response analyzer (FRA), 103–106
FujiFilm Dimatix printer, 183

E
Economic and business perspectives, 277–291
Effective mass, 18
Electrochemical anodization, 307
Electrochemical nanostructuring, 135–137
Electrochemical scanning light beam analysis, 181–187
Electrochemical synthesis, 194
Electrolyte, 233
conductivities, 79
fluctuating energy levels, 45
voltage drop, 77
Electromechanical shutters, 93
Electrolyte, 233
conductivities, 79
fluctuating energy levels, 45
voltage drop, 77
Electromechanical shutters, 93
Global population, 3
Global warming, 4
Gouy layer, 39
Green Economies, 205
Guest-host nanostructures, 9

G
GaAs, 240
GaInP₂, 240
Gärtner equation, 49
Gärtner model, 296, 299
Gauss’ law, 32
General design challenges, 245–247
Gibbs free energy, 14
Global population, 3
Global warming, 4
Gouy layer, 39
Green Economies, 205
Guest-host nanostructures, 9
Helmholtz capacitance, 111
Helmholtz layer, 37
Hematite, 187. See also Iron oxide
H₂/H⁺ potential vs. vacuum, 39
High pass filters, 93
High surface area morphology, 196
High-throughput screening, 168–170
Host-scaffold/guest-absorber approach, 143
Hybrid photoelectrodes, 256–259
Hydrogen, 6, 173
Hydrogen cost targets, 282
Hydrogen Economy, 208
Hydrogen evolution efficiency, 53
Hydrogen storage, 7, 288

H
Helmholtz capacitance, 111
Helmholtz layer, 37
Hematite, 187. See also Iron oxide
H₂/H⁺ potential vs. vacuum, 39
High pass filters, 93
High surface area morphology, 196
High-throughput screening, 168–170
Host-scaffold/guest-absorber approach, 143
Hybrid photoelectrodes, 256–259
Hydrogen, 6, 173
Hydrogen cost targets, 282
Hydrogen Economy, 208
Hydrogen evolution efficiency, 53
Hydrogen storage, 7, 288

I
Impedance measurements
instrumental limitations, 107–109
measurement parameters, 106–107
quasi-reference electrode, 109
reference electrode, 109
Incident photon-to-current conversion efficiency (IPCE), 54
Inkjet printing, 181–187
Intellectual property, 289
Intercalation, 52
Interface design, 245
Interfacial electron transfer, 45–47
International Panel on Climate Change, 4
Inversion layer, 34
Ionic bonding, 15
Ionic charge compensators, 180
Ionic compensation, 23
IPCE. See Incident photon-to-current conversion efficiency (IPCE)
IrO₂, 10
Iron oxide, 304
absorption coefficient, 125
charge carrier transport, 126–129
combinatorial screening, 187
crystal structure, 123
density of states, 125
electron transport, 127
excited state lifetime, 126
flatband potential, 129
hole diffusion length, 129
hole mobility, 306
magnetic properties, 124
morphology control, 131
nanotubes, 306
optoelectronic characteristics, 124–126
structural and material properties, 122–123
ITO, 74

J
Jarrett’s model, 50
Junction under illumination, 225–227

L
Lenses, 92
Long pass filters, 93
Loss minimization, 232–234

M
Material parameters, 287–288
Maximum power point, 243
Metal nanoparticle, 309
Metal nanostructures
scattering, 311
Metal organic decomposition (MOD), 163
Metals, functionalities in photoelectrodes, 180
Methane, 6
Methanol, 6
Midgap states, 21
Minority carrier diffusion length, 56
Mixed metal oxide photocatalysts, 157–170
Mixed metal oxide photoelectrodes, 157–170
Mixed metal oxide semiconductors, 158–162, 179
Molecular orbital (MO) diagram, 15, 16
Molecular sensitizers, 181
Monochromatic light sources, 90
Monochromator, linear dispersion, 92
Monolithic PEC device, 9
Morphology control, 131
Mott-Schottky equation, 43, 101
Multicomponent materials, 179
Multi-junction approaches, 205–267
Multi-junction design considerations, 248–249
Multi-junction device, 247
Multi-junction device schemes, 253–263
Multi-junction efficiency limits, 249–253
Multijunction III–V semiconductor technology, 206
Multi-junction stacks, 236–237
Multiplexed photoelectrochemical screening system, 195
Nanocrystalline electrodes, 159
Nanomaterials, 294–308
Nanophotonics, 308–312
Nanostructured $\tau$–$\beta$-$\text{Fe}_2\text{O}_3$, 121–150
Nanostructured electrodes, 58–60, 112
Nanostructured morphology, 197
Nanostructuring method, 145–146
Nanotube arrays, 302–308
Nanowire array photoanode, 59
Nanowire arrays, 134–135
Natural gas, 3
Near-field absorption, 310
Near-field absorption enhancement, 312

OH groups, 30
Oil, 3
Oil consumption, 210
Open-circuit voltage, 242
Optical absorption cross section, 26
Optical considerations, 237–241
Optical density, 54
Optical transition, 17
Optimum stoichiometry, 183
Overpotential, 9
Oxidation reaction, 14
Oxygen evolution overpotential, 147–148
Oxygen evolution reaction, 9

Peak oil, 3, 279
PEC configurations, comparison, 177
PEC costs for hydrogen, 282–283
PEC water splitting losses, 221–222
PEM electrolyzers, 283
pH bias, 174
Phosphate buffer, 79
Photocatalyst reactor, 215
Photocatalyst Z-schemes, 262–263
Photocorrosion, 51–53, 178
Photocurrent action spectra, 98–99
Photocurrent transients, 97–98
Photocurrent–voltage characteristics, 49–50
Photodiodes. See Calibrated photodiodes
Photo electrochemical cell, 14
Cappuccino cell (EPFL), 73
cell design, 71
conducting substrate, 74
sealants, 73
TU Delft design, 72
Photo-electrochemical hydrogen plants, 212–217
Photo electrochemical (PEC) measurements
AM1.5 performance, 100
current–voltage curve, 95
front-and back-side illumination, 96–97
photocurrent onset potential, 95
simulated sunlight, 80–81
water splitting efficiencies, 95–96
wavelength-dependent, 81–82
Photoelectrochemical pn-junction diode, 305
Photoelectrode criteria, 293
Photoelectrode junction formation, 225–227
Photo-electrodes, requirements, 9, 55
Photoelectrode system, 217
Photoelectrode system with moderate concentration, 217
Photo-reduction, 51
Photovoltaic–electrolysis systems, 9
Photovoltaic multi-junctions, 246–248
Plasmon, 310
Plasmons, 309–312
Point of zero charge (PZC), 36
Poisson’s law, 30
Polaron, 294
Polaron mobility, 295
Polaron size, 295
Polaron transport, 296
Porous thin films, 132–134
Potential distribution, 30–34
Potentiostat
diagram, 83
specifications., 84
Potentiostatic anodization, 135
Printers, 182
Printing template, 185
Protonation and deprotonation reactions, 36
PV-electrolysis, 283
PV-electrolysis approach, 214
Quantum efficiency, 99–100, 239
external, 54
internal, 54
Quantum size effects, 59
Quasi-Fermi levels, 48, 228
Quasi-reference electrode, 109–110
Real estate issues, 285–286
Recombination, 98
Recombination in the space charge layer, 296
Reduction reaction, 13
Reference electrode, 40
overview of different types, 76
Reichman equation, 49
Reorganization energy, 45
Requirements for porous systems, 160
Reverse water–gas shift reaction, 8
RHE scale, 75
Routes to solar hydrogen, 8
Routes to solar water-splitting, 211
Route to market, 290
Ruthenium dye, 10
Rutile TiO2, 51

S
Sacrificial reagents, 160, 177, 179
Scale-up problems, 285
Scanning laser screening systems, 182
Shallow dopant, 20
SHArK Project, 193
Shockley surface states, 44
Short-circuit current-density, 242
Shutters. See Electromechanical shutters
Si doped Fe2O3, 141
Silver nanoparticles, 310
Silver nanowires, 311
Single bed photocatalyst particle suspension system, 216
Single-bed photocatalyst system, 217
Single junction efficiency limits, 234–235
Single junction examples, 235
Single junction PEC semiconductor devices, 225–235
Solar-cell–electrolysis, 157
Solar cell market, 6
Solar cells, 5
Solar concentration and tracking, 289
Solar conversion pathways, 211–212
Solar photocurrent, 54
Solar resource, 209–211
Solar simulator class, 88
Solar spectrum considerations, 286
Solar-to-hydrogen conversion efficiency, 223–225
Solar-to-hydrogen efficiency, 53
Sol gel routes, 195
Soluble metal oxide precursors, 182
Space charge layer
charge density, 32
potential drop, 33
total charge, 33

Space charge layer width, 297
Space charge width, 33
Spray pyrolysis, 138–140
Storage of solar energy, 281–282
Surface catalysis, 147–148
Surface hydroxylation, 36–37
Surface plasmon, 310
Surface states, 110
Sustainable energy sources, 4
wind, 4
System costs, 289

T
Tamm states, 44
Tandem cell, 146–147
Tandem device, 62
Tandem photoelectrodes, 177, 255–256
Tandem photoelectrolysis system, 176
TaON, 162
Techno-economic analyses, 217–219
TiO2, 303
Transparent conducting oxides, 74
Transparent conducting substrates, materials parameters, 74
Trap-mediated recombination, 300
Triple junction hybrid photoelectrodes, 259–261
Tunneling, 46
Two-bed photocatalyst, 216
Two electrode PEC system, 214

U
Unpinning of the band edges, 44
US department of energy, 9
Useable photo-potential, 229–231
Useable photo-voltage, 241

V
Vacuum level, 38
Value of hydrogen, 208–209

W
Water, 6
Water oxidation, reaction mechanisms, 50
Water-splitting reactions, 220–221
WO3, 160, 260

Z
Zinc oxide, 22