4

Cell-Balancing Techniques: Theory and Implementation

4.1 Introduction

In the safety chapter we briefly discussed the issue that when multiple cells are connected in series, the cell voltage is not always equal to the pack voltage divided by the number of cells. How does this happen? This chapter explores that question in detail, but the first question to answer is this: Why do we care?

The first reason is safety. Remember, when lithium ion cell voltage exceeds 4.2V by a few hundred millivolts, it can undergo thermal runaway, melting the battery pack and device it is powering. It can even blow up as a big ball of fire. Although a well-designed pack has an overvoltage protection circuit that will prevent such an event (usually even two independent circuits!), it is better not to tempt fate by triggering this protection unnecessarily.

The second reason is longevity. If the maximal recommended charging voltage is exceeded even a little, it will cause very accelerated degradation. Just increasing the charging voltage from 4.2 to 4.25V causes the degradation rate to increase by 30%. For this reason, the misbehaving cell that has higher than its due share of voltage will degrade faster.

The third reason is incomplete charging of the pack. Let's assume the protector circuit does its job and that charging stops when just one cell gets close to unsafe conditions. Now we have successfully prevented thermal runaway, but all of the other cells now have lower voltages and are not fully charged. If we look at the pack voltage, it will be much less than 4.2V multiplied by the

number of cells. Less pack voltage means less pack energy. (It also usually means less available capacity, as we will see later.)

The fourth reason is incomplete use of pack energy. Let's consider another situation. Instead of having too high a voltage, one cell could have too low a voltage compared to others when the pack is close to the end of discharge. A pack protector will prevent overdischarge (which would damage the cell) by stopping the discharge of the whole pack when one cell voltage goes below the cell undervoltage threshold (usually around 2.7V for a $LiCoO_2$-based cell). This means that all other cells are still at higher voltages and have energy left. The pack still has energy, but the device can no longer be used because of one misbehaving cell.

Now that we have established that cell voltage differences are harmful enough to take action to remove them, let's look at the causes of these voltage differences. The first thing to understand is that a voltage difference is not in itself an imbalance, but a manifestation of the differences in state of charge (SOC) of the cells if no current is flowing, and of the cell resistance differences if current is flowing. If we try to instantly eliminate the voltage differences themselves (e.g., the "effect") without eliminating their cause, we will potentially do more harm than good, while wasting hardware resources (cost, size) by overengineering the balancing circuit to provide huge currents that will be required for such an instant result, and wasting energy by unnecessarily passing currents back and forth. Unfortunately, this is exactly what happens in some commonly used balancing schemes that have been designed without an understanding of the underlying mechanisms of the imbalance. The plan is that after reading this chapter you will never design or use such an inefficient system because you will be endowed with a perfectly clear idea of what an imbalance is and how it can be eliminated in the most theoretically efficient way.

4.2   Types of Battery Cell Imbalance That Affect the Charge/Discharge Voltage

4.2.1   State-of-Charge (SOC) Imbalance

A SOC difference is the only cause for cell voltage differences if no current is flowing, known as open circuit voltage (OCV). Indeed, there is a simple correlation between SOC and voltage for any battery chemistry in the form of OCV = f(SOC, T), where SOC is the state of charge and T is temperature. The form of the function is different depending on the chemistries, but in general it is clear that for a given dSOC you get some difference in voltage, dOCV. What could be causing these differences in cell SOCs? Let's look at a few possible reasons.

First we consider inaccuracies in the voltage measurement of cell formation cyclers. Most Li-ion cells are cycled after assembly to form a passivating layer on the anode and to detect abnormal cells. At the end of the cycling, all cells should end up in the same state of charge (as indicated by the same voltage). However, cycling equipment is not perfect; there are some channel-to-channel variations that result in cell SOC differences. To reduce these, of course, the cyclers themselves have to be kept well calibrated, but also, after cycling, cells need to go be graded, a process in which the cells are grouped based on close voltage, usually within 2 mV from each other. Now, 2 mV may not sound like much, but keep in mind that cells are stored and delivered in 50% SOC. It happens to be close to the flattest portion of the voltage curve, where a 1-mV difference roughly corresponds to a 1% SOC difference. This same 1% difference in SOC will result in up to a 10-mV difference by the end of charge and a 100- to 500-mV difference by the end of discharge (depending on how deeply the pack is discharged) because the voltage/SOC curve is much steeper in these areas. Note that while the percentage of SOC imbalance remains constant during the entire discharge period, voltage differences among the cells vary with SOC because $dV/dSOC$ varies with SOC. Figure 4.1 shows OCV differences among cells at a constant SOC imbalance but at different states of charge.

This figure shows the dependency for a Li-ion cell. The actual shape of the curve will vary with chemistries, but the concept that SOC differences remain constant regardless of SOC is still valid. Some chemistries, such as a lead-acid chemistry, experience large and almost linear changes of voltage with SOC so it is quite easy to estimate how much $dSOC$ is between the cell for a given dV. Other chemistries, such as $LiFePO_4$, have an almost completely flat voltage profile, so even differences in SOC between the cells as large as 5% to 10% are not noticeable in the cell voltage when it is in the midrange of SOC but can nevertheless cause drastic voltage deviations close to the end of charge and discharge, causing protection electronics to trip and to shut down the charging process before the pack has a change to be fully charged. It is also more difficult to grade such cells by voltage, which makes the need for in-system balance management more critical for this chemistry. But such a system clearly cannot be based just on voltage in this case. Later we will look at systems that work for "flat" chemistries.

The second important reason for SOC differences between the cells in a pack are differences in self-discharge rates between the cells. The self-discharge rate is strongly dependent on temperature. It approximately doubles with every 10°C increase from room temperature. System design does not always take into consideration the need to heat the pack evenly, and places various heat-generating components such as the application processor, backlight, and memory in such a way that they can fit most of the components into the smallest
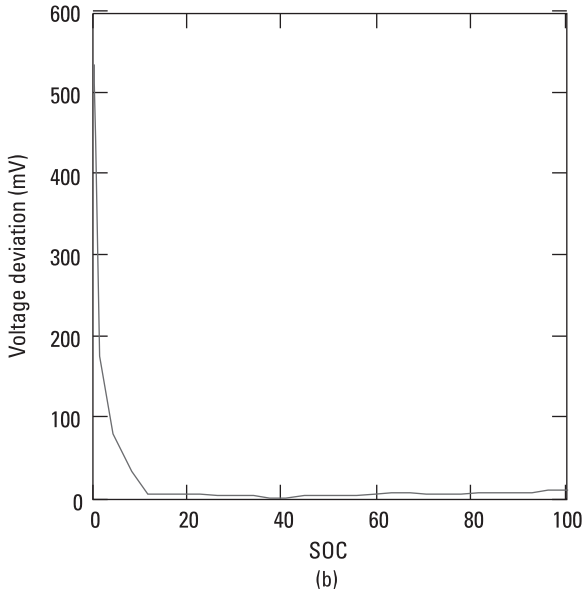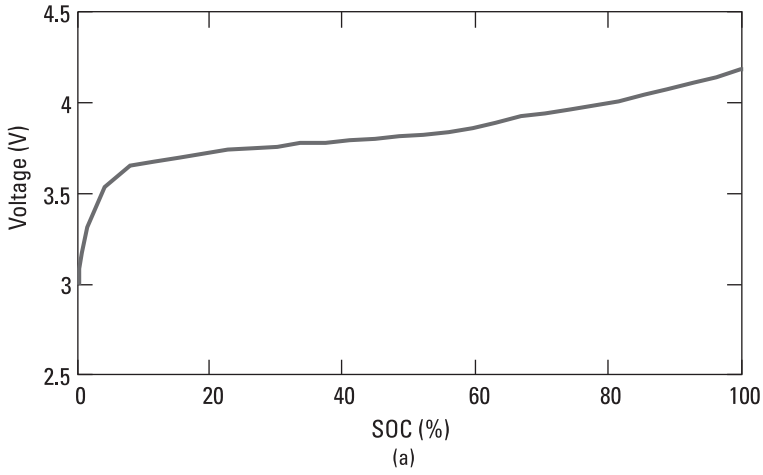
Figure 4.1   (a) OCV dependence on SOC. (b) OCV differences at different states of charge between two cells with a SOC imbalance of 1%.

space, rather than ensuring that cells in a pack will have the same temperatures. This will cause one cell that is hotter to "leak" more charge than a cell that is cooler. This cell's SOC will gradually decrease, because the charger puts the same amount of coulombs into each serially connected cell, but some coulombs in the hot cell get internally short-circuited due to self-discharge and do not contribute to increasing SOC. This process is quite slow, because even at the highest temperature of 60°C, the self-discharge of a Li-ion battery is only about

50% in a year. But over time, the difference in SOC can become significant because the effect is accumulative (the same cell will usually be hotter than the others). Battery pack design could include a heat spreader between the system and the pack, and a good cooling surface on the other side of the cells to keep their temperature and self-discharge in general to the minimum. That also helps to decrease cell degradation, which also happens to be strongly accelerated by temperature.

A SOC imbalance can be also caused by uneven leakage to the battery pack circuit from different cells. It is easier to power pack electronics from the low voltage of just one cell (usually the one closest to the pack ground), because cheaper low-voltage ICs can be used. It looks like a nice shortcut to take, because the electronics would only consume some hundreds of microamps, which is negligent compared to multi-ampere-hour size cells. Unfortunately the effect is accumulative. Over a long period of time, the difference in the amount of charge removed from the lowest cell will increase and eventually reach substantial numbers. For example, if we have three 2,200-mAh cells (Q_{max}), and discharge one by 100 mAh (Q_1), the second by 100 mAh, and the third by 100mAh of actual load + 100 mAh accumulated from 1 month of low current due to powering some circuit, the first and second cells' chemical state of charge will be ($Q_{max} - Q_1$)/Q_{max} = 95.4%, but the third cell will be at 91%. So we can say that cell 3 is imbalanced by 4.4%. This in turn will result in a different open circuit voltage for cell 3 compared to cells 1 and 2, because the OCV is in direct correlation with the chemical state of charge. This problem can be resolved by good pack design, which powers pack electronics only from the entire pack voltage. It also assures that all of the connections to each serial cell (for example, those used for voltage measuremetns) are high-ohmic. Typically, these connections are ADC inputs with impedance in mega-ohm ranges, and do not draw more than a microamp or two. In addition, this impedance is the same for all cells, so even the tiniest current drawn does not cause an imbalance. All major safety and gauging ICs from Texas Instruments, Intersil, or Maxim are powered from the pack voltage and avoid this issue.

4.2.2   Total Capacity Differences

Sometimes cells have the same voltage at the end of a charge and, hence, appear to be perfectly balanced (and correspondingly have the same SOC), only to show a very large deviation of voltage at the end of discharge. Inversely, cells having an equal SOC and voltage in the discharged state can show large differences at the end of charge. What is the reason for this mystery? It happens to be caused by differences in cell capacities. Indeed, if you have two cups of different heights, they will both be equally empty (e.g., state of charge zero and perfectly balanced) when there is no water in them. But once you pour an equal amount

of water into both cups, the taller will be half full, but the shorter will be completely full (e.g., states of charge 50% and 100%, respectively).

The dilemma in such a situation is that you cannot just balance cells with different capacities once and have them stay balanced during charge/discharge (as was the case with SOC balancing for the cells of the same capacity when not given enough time for self-discharge). We have four choices:

1. Balance the cells on top (in the fully charged state) and let them diverge at the bottom. This will be repeatable—they will always stay balanced on the top after that without any further action.

2. Balance the cells on the bottom (in the discharged state) and let them diverge at the top. Again, once the initial balancing is done, it will be a persistent state.

3. Completely rebalance the cells during each discharge, for example, extract energy from higher capacity cells on the way down so that by the time it reaches the bottom all cells have the same energy. This would also require a complete rebalancing during each charge, so that the moment the energy from smaller capacity cell becomes full is delayed until other, larger capacity cells become full. Going back to our analogy with the cups, this would be equivalent to leaking some water from the jar past the shorter cup when poring it so that it will be at the same level as the taller cup at the end.

4. Finally, there is a variation of choice 3, where we are not just throwing away the energy that we take from the larger cell on the way down, but are actually passing it to the lower capacity cells. Inversely, on the way up, energy from smaller capacity cells is not just dissipated, but forced into larger capacity cells. This way, in our example with the cups, we will not end up with two cups at just the 50% level (e.g., matching larger cup), but with both cups at some intermediate higher level, say 75%, because we did not spill any water on the table. It also helps with reducing the subsequent cleaning.

The best choice appears to be obvious (choice 4), but in reality it is not so simple. Because nothing in life is free, passing energy from one cell to another requires an actual hardware implementation (discussed below). Let's just say that it has inductors (or capacitors) and high-power FETs so it is about as large as all of the safety circuits and gauging circuits taken together, so you would likely double the size and cost of your overall battery management solution. Before making a heavy investment in additional hardware, we have to consider how much benefit will we get. That will depend on several factors:

1. How much cell capacity imbalance is there? Little imbalance, little benefit.

2. How large is the cell? Sometimes it is so large and expensive that the additional cost of the balancer would be justifiable to get all of the energy from such a large system.

3. Is heat dissipation (e.g., spilled water on the table) a problem in this particular design because batteries are large and space is confined?

Actually, for many laptop designs it turns out that capacity differences after factory grading can be reduced to 1% to 2%. So in many cases it is cheaper to use a 1% to 2% larger battery rather than investing in doubling the size and cost of hardware.

Now, if we are not going to actually transfer capacity from cell to cell, what is the next best option? Well, closer observation shows that option 3 requires a lot of "spilling of water" both on the way up and on the way down, and it has to happen for every single cycle. Heat dissipation due to damping of additional energy into the environment can be undesirable because it requires installation of additional cooling capacity. But most importantly, what do we gain if we completely rebalance the cells on the way down? All cells will reach 0% at the same time but this will be the same time as the lowest capacity cell would reach it without any balancing. A battery pack protector would have shut down the discharge of the pack at the same time anyway, because the lowest capacity cell will trigger an undervoltage condition. So we released all of this heat from higher capacity cells (at the cost of some additional hardware!) and improved run-time by … zero seconds.

How about the charge direction? There is some benefit to keeping cells from exceeding the recommended charging voltage, so if we can prevent the lower capacity cells from going all the way to the overvoltage threshold (which can be as high as 4.3V for 4.2V normal charging voltage cells), we could improve their health and stop the vicious cycle where the pack gets overcharged, degrades more, ends up with even lower capacity, overcharges more, degrades more …, repeat from the beginning. But do we have to do it every cycle? Actually, if we give up on the idea of rebalancing the cells on the way down (which, as we have seen, has no merit) then we do not have to rebalance them on the way up either! We can just balance it once, on the top, and it will stay this way without any further action! Which brings us to choice 1, which is indeed the most popular choice in portable electronics. Choice 1 is often called *top balancing* and choice 2 is often called *bottom balancing*. To summarize, bottom balancing has no merit in systems that utilize energy dissipative balancing, whereas top balancing helps preventing excessive degradation of a lower capacity cell.

But wait, do you have this question on your mind: If we need to balance only once, why do we even need the balancing circuit in the pack? We could

just balance cells on the top to correct the capacity differences during assembly (e.g. "gross balancing") and be done with it, all the while saving money and space. Well, this statement indeed makes sense. We really do not need an expensive and spacious high-current balancer designed to provide such high currents that would allow keeping up with high rate external loads and compensating cell capacity differences in real time. We still need a small lower power balancer to continuously counteract *other causes of SOC imbalance* that are themselves continuous and accumulative in their nature—the differences in self-discharge rate of cells due to their different temperature, for example.

In addition, the cells degrade at a slightly different rate, so cells that had the same capacity initially will gradually develop some "individuality." So this additional discrepancy has to be balanced out, as it develops, in the pack itself. However, both the self-discharge rate and degradation rate are extremely slow and for that reason only balancing current in the milliamp range is needed to keep it in check. Such low current needed for *maintenance balancing* can be easily provided by integrated circuits that are already there, such as protector circuits, without any added cost or size. We look at some specific examples shortly.

4.2.3   Impedance Differences

In addition to SOC differences, another cause of voltage discrepancies among cells is the cell impedance difference. However, this has an effect only if current is flowing. How large is the effect? It depends on cell impedance differences and the current. Internal impedance differences among the cells can be expected in the ~15% range in the same production batch as can be seen form the example test data shown in Figure 4.2(a).

Impedance imbalances do not cause differences in fully relaxed OCV, when any effect from current flow is already dissipated. However, they will
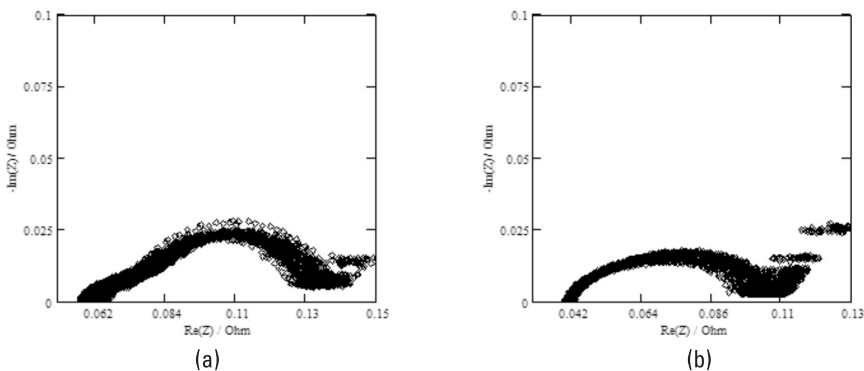

Figure 4.2   Impedance spectra differences between 50 cells in one batch for manufacturer (a) and 50 cells for manufacturer (b). Data shown range from 1 kHz (left) to 10 mHz (right).

cause differences in cell voltage during discharge. For steady-state current flow, cell voltage can be approximated as $V = OCV + I * R$, where R is the low-frequency portion of cell internal impedance (right side of the graph in Figure 4.2). If the current is negative (discharge), the voltage will be lower for a cell with higher R. If the current is positive (charge), the voltage will be higher for a cell with higher R.

As can be seen from Figure 4.2, high-frequency impedance at 1 kHz (left side of the graph) is well matched for manufacturer (b) and not as well matched for manufacturer (a). However, low-frequency impedance (the one that will actually matter for continuous discharge) is equally badly matched for both. The reason for this is that most cell makers have access to simple 1-kHz impedance meters that allow them to grade the cell based on high-frequency impedance. High-frequency (1-kHz) measurement is very fast and allows for the detection of massive failures such as a short circuit or current collector disconnect. However, it is not very useful to observe the whole range of electrochemical properties of the battery related to actual charge storage. Low-frequency impedance would be more useful for preventing a voltage imbalance, but its measurement takes at least 10 seconds, so it is very rarely used in cell production (but could be used by a pack maker for improving the cell matching).

No balancing algorithm can eliminate the resistance differences; they are a permanent property of a battery pack once assembled (that is why preassembly grading is beneficial). In fact, the imbalance can increase with aging because different cells' impedance is likely to change at a slightly different rate. But impedance differences need to be considered in any balancing scheme, especially one based on voltage, because they can significantly distort attempts to balance what we can and should balance; namely, the SOC. Note in Figure 4.3 that for the absolute majority of discharges (from 10% to 100% SOC) the distortion of voltage that is caused by the impedance deviation is larger than that caused by a SOC imbalance.

By looking at voltage alone we cannot distinguish which part of the cell deviation is due to a SOC difference and which part is due to an impedance difference. Both parts can shift voltage in the opposite direction! If we do not know about this and just "assume" that all voltage shift is due to a SOC difference, we might be tempted to correct it by bypassing some charge through the cell with the higher voltage that "appears" to have a higher SOC.

However, if most of the difference is caused by an impedance imbalance (as is commonly the case), bypassing more current through this cell will result in the opposite effect—it will increase the SOC difference from other cells to a larger value than it would be without balancing. As a result, the open circuit voltage of this cell at the end of charge will be different from the other cells and can reach high levels, potentially causing the safety circuit to trip.
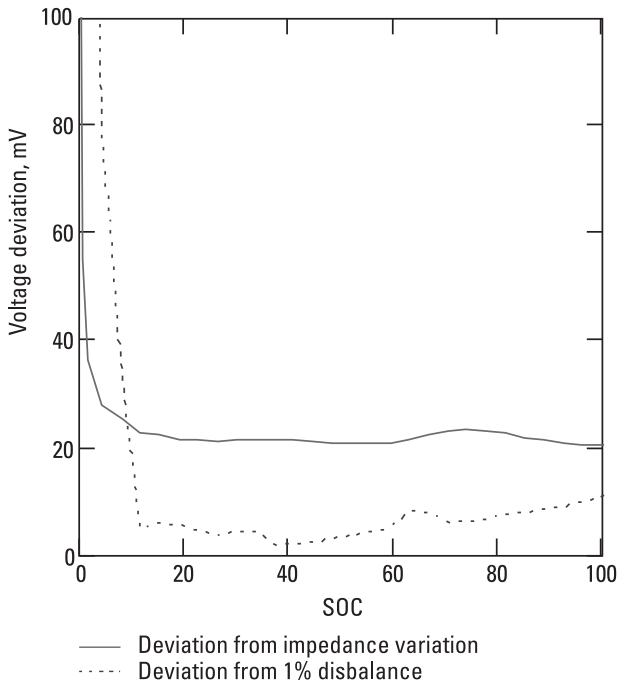
Figure 4.3   *Solid line:* voltage differences between two cells with 15% impedance imbalance at C/2 discharge rates. *Dotted line:* difference between the cells with 1% SOC imbalance for comparison.

If, for example, we are using a simple balancing scheme when a bypass FET allows us to turn on a load parallel to any cell and the bypass FET is turned on based on voltage during charge, it can cause an actual increase of the imbalance through bypassing the cell with the higher impedance. At the end of the charge, the IR rise becomes insignificant because of current decrease, so that the FET switches on at the other cell. However, it happens too late so at the end of charge this procedure results in a higher SOC and higher voltage for low-impedance cells. Eventually it will lead to increased cell degradation. This problem can be reduced if cell balancing only switches on near the end of the charge when the current is reduced and so the I*R drop has a smaller effect on battery voltage.

It is more difficult to fight the IR effect during discharge, because there are usually no predictable periods of low current except for some rare application. Fortunately top balancing (which we earlier found to be the only one useful for "dissipative" implementation) is mostly done only during charging.

There are periods of low current (e.g., in the case of inactivity, rest) that potentially could be used for balancing without IR effect, but it would be still beneficial only if balancing could be done that would ensure the cells were equal

on top (at the end of charge), which requires "predictive balancing" rather than "reactive balancing" unless the rest is happening in the fully charged state. See the later discussion on balancing algorithms for more details.

Note that distorting due to impedance differences is inherent to any voltage-controlled balancing method (such as for inductive energy redistribution) and not just to simple bypass balancing. In fact, in energy redistribution cases, the effect can be even more dramatic because energy will be passed back and forth with the high currents that such balancers are typically capable of, which can cause overheating, loss of efficiency, and even complete instability of the control circuit.

Another effect of battery impedance on voltage imbalance exists regardless of any impedance differences. This effect just amplifies voltage differences due to a SOC imbalance. Again, to explain the effect we can model the voltage under steady-state load as $V = OCV(SOC) + I * R(SOC)$ (considering that discharge current is negative). Because function $R(SOC)$ is rapidly increasing its value as SOC approaches zero, the voltage differences between the cells with fixed SOC imbalance increase in highly discharge states, as shown in Figure 4.4. This gives the impression that there is an increased need for balancing near the end of discharge. However, if the SOC imbalance is removed during other stages of discharge and is absent by the time low SOC is reached, the increased
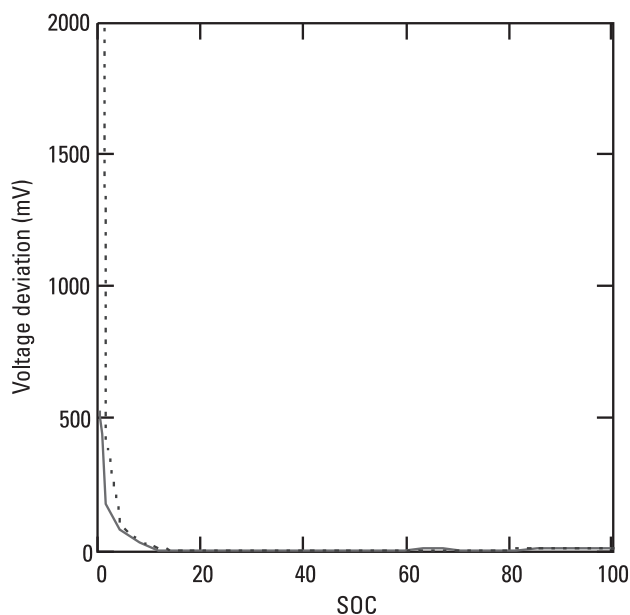

Figure 4.4 Voltage differences under C/2 load at different states of charge among cells with a 1% SOC imbalance. *Solid line:* Differences for OCV case for comparison.

voltage differences near the end of discharge will be eliminated without need for high bypass currents.

4.3   Effect of Imbalancing on Performance

4.3.1   Premature Cell Degradation Through Exposure to Overvoltage

Now that we have a clear understanding of the underlying mechanisms of voltage differences, we can evaluate in more detail different issues that some of these mechanisms can cause. Impedance differences will cause cell voltage differences during charge and discharge, but because this difference in voltage is purely ohmic in origin (to large extent), just like IR drop/rise across a resistor, it does not cause accelerated cell degradation as such and will disappear once the current is turned off. So although there is nothing we can do about this impedance difference, there is also no need to do anything. On the other hand, in the cases of a SOC or total capacity imbalance, the cell with the higher resulting SOC is exposed to higher chemical potentials that will cause accelerated degradation. For example, what happens if one cell has less capacity than the other three serially connected cells in the pack, if they all start at the same state of charge? CC/CV charging will bring the pack to $4.2 \times 4 = 16.8V$ (typical). However, individual cell voltages will not be equal. As you can see in Figure 4.5, the "low-capacity" cell will have a much higher voltage than the remaining cells, while the normal capacity cells will have a lower voltage than is achieved in normal charging. As the cell is exposed to higher potential, it will degrade more, thus
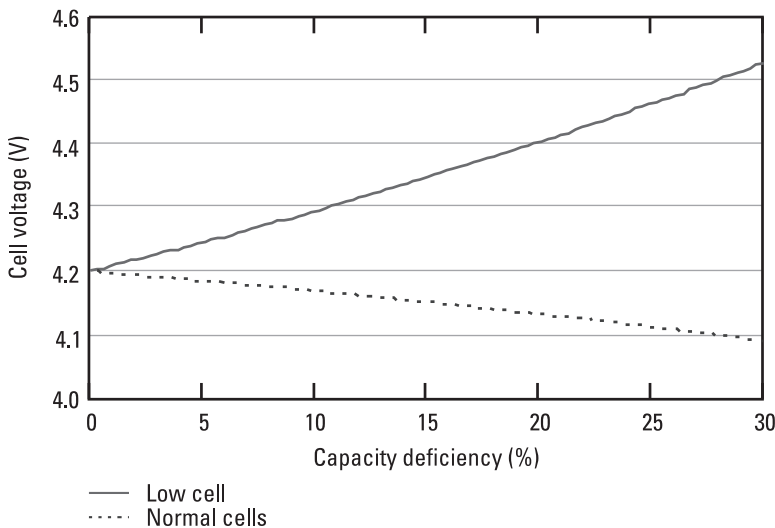

Figure 4.5   Individual cell voltage versus capacity deficiency from nominal.

increasing the capacity deficiency, which will move the pack to the right on the graph in Figure 4.5. Eventually, when the lower cell reaches a total capacity deficiency above 10%, its cell voltage rises into the dangerous area above 4.3V, which will result in extreme degradation of this cell or even become a safety concern.

Note that not all battery chemistries are equally affected by cell voltage imbalances at the end of charge. While the Li-ion chemistry is especially vulnerable because of its ability to store almost 100% of all energy delivered with negligible self-discharge in its operational voltage range, lead-acid, NiMH and NiCd chemistries are relatively tolerant to overcharge because they can respond to increased voltage by internal shuttle reactions that are equivalent to a chemical short circuit inside the cell. For example, in a NiMH battery oxygen and hydrogen generated after the end of charge recombine inside the cell, building water. This causes extensive heating because all the energy of the charger is converted to heat rather than stored, which is undesirable but at least it does not cause thermal runaway. Still, overcharge at high rates does cause increased pressure inside the cell and will accelerate cell degradation and can even create a chance for explosion or venting. The need for cell balancing has to be evaluated in conjunction with rate capability, cooling, and other properties of the charging system.

4.3.2 Safety Hazards Resulting from Overcharged Cells

Li-ion batteries have very high electric energy concentrated in a small volume. While the possibility of its release via a short circuit can be prevented by appropriate mechanical protections, the coexistence of highly reactive chemicals in proximity makes this battery inherently dangerous. Overcharging and overheating of the battery cause the active components to react with electrolyte and with each other, ultimately causing an explosion and fire. Thermal runaway can be caused merely by overcharging a single cell to voltages above 4.35V. Other cells of the pack will also join the explosive chain reaction if one cell is compromised. That is why continuous cell balancing should prevent any cells from getting anywhere near the dangerous voltage territory, and a safety protection circuit should terminate the charge if this somehow happens.

4.3.3 Early Charge Termination Resulting in Reduced Capacity

Additional safeguards present in the battery pack primary protectors (such as bq30z55) and the independent secondary protectors can help alleviate safety issue. The protector and gas-gauging IC will terminate charging if one of the cell's voltages exceeds the programmable cell overvoltage threshold (default 4.35V). An overvoltage protector will terminate charging and prevent an unsafe condition from occurring, but at the same time it will keep the whole pack un-

dercharged (since other cells are now in a much lower state of charge). Overall, the pack loses energy as the result because the higher voltage in the steeply rising area close to a fully charged state does not compensate for the large capacity loss in much more flat areas where the mAh/volt are much higher.

Because the effects of cell degradation caused by imbalance are auto-accelerating, preventing such a vicious cycle allows us to extend battery pack life significantly and to provide a longer run time despite some initial capacity imbalance.

4.3.4 Early Discharge Termination

So far we have mostly discussed cell degradation if overcharged. This is the greatest safety concern and also the most common degradation mechanism. However, the cell will also be damaged if it is severely overdischarged. In fact, if cell voltage goes below 2V, the actual dissolution of the Cu-current collector will occur. This is deadly to the cell if the process is allowed to continue to a significant extent (no current collector, no charge/discharge). For this reason any device using a Li-ion battery absolutely has to have a discharge termination based on voltage. Some naïve attempts to introduce a consumer-replaceable Li-ion battery, for example as a drop-in replacement for coin cells, have so far faltered for the very reason that primary cells that are being replaced did not require any undervoltage shutdown (they just die, being primary) and might keep draining substantial current regardless of battery voltage even when no useful operation is taking place. If a Li-ion cell is placed into a device that was not designed for it, the cell will be overdischarged below 2V and die just like the primary cell that it replaces.

This is mostly not an issue for devices designed to use Li-ion cells, because they will have a full-off state when the pack voltage goes below a safe threshold. In addition, the battery pack itself will often have a battery management system. To prevent overdischarge of cells and resulting damage, battery managements system will terminate the discharge process if any of the cells reach the pre-determined low-voltage threshold. The cell-based termination voltage is usually set to a lower value than the pack-based threshold divided by number of series cells, so that the difference can allow for a small imbalance. For Li-ion batteries, the threshold varies from 2.7 to 2.2V depending on the typical discharge rate. However, a larger imbalance will cause the overall pack to terminate when a cell with lowest capacity or SOC will reach cell undervoltage while other cells still have energy left.

Redistributing the energy from higher cells to lower cells during the end of discharge phase can increase a battery's useful discharge time. Note that if "bottom balancing" is being used, it needs to be a redistribution of energy, and not just a bleeding out of extra charge; otherwise, it would just dissipate heat

and not increase the run time. To be effective, an inductive or capacitive energy redistribution circuit with high efficiency (usually above 80%) is required. If the control algorithm reacts only to voltage differences that become noticeable only very close to the end of discharge, it would also require a high-rate bypass capability to keep up with the high discharge current. Such circuits are expensive to implement in redistribution balancing circuits, and larger inductors and FETs use up space that is at premium in portable devices. An approach that utilizes the hardware more effectively would be to gradually redistribute any existing SOC imbalance during the entire charge/discharge process, not just when it results in acute voltage differences (at the end of discharge); this is known as *predictive balancing*. This, of course, requires the ability to determine how much charge needs to be bypassed somehow without relying only on voltage. How to accomplish this is discussed in the balancing algorithms section.

4.4 Hardware Implementation of Balancing

4.4.1 Current Bypass

One simple implementation of cell balancing uses a MOSFET in parallel with each cell and controlled by a comparator output for simple voltage-based algorithms that turn on the bypass FETs during the onset of voltage differences, or controlled by a microcontroller for more complex and effective algorithms that can work continuously regardless of variations in the voltage. A general setup is shown in Figure 4.6.

The main choice here is to use MOSFETs that are integrated in the balancing controller IC and typically have bypass currents from 9 to 2mA (depending
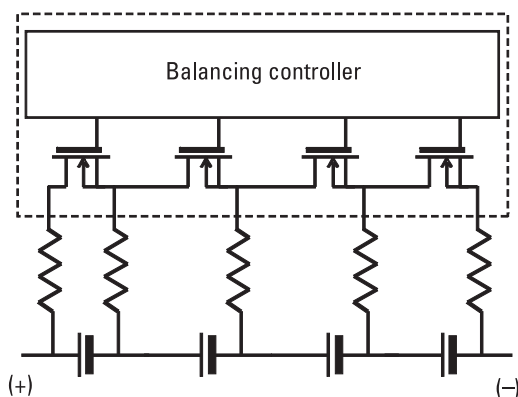

Figure 4.6 Cell-balancing setup using bypass FETs.

on the choice of the external resistors), or to use external FETs with bypass capability that can be freely tailored to particular application needs.

In Li-ion batteries that have a very low self-discharge capability and, therefore, an accumulative imbalance per cycle of usually less than 0.1%, the bypass current of internal FETs is sufficient to keep the pack continuously balanced. In other chemistries where self-discharge rates are much higher and, therefore, differences in the self-discharge rates among the cells result in higher SOC differences per cycle, higher rates might be needed. Some balancing circuits have separate pins for voltage measurement and charge bypass; however, this is not common in portable devices because the larger number of pins increases the size of the device, which is a disadvantage in constrained spaces. The issue with balancing current interfering with voltage measurement is usually addressed by the firmware turning balancing FETs on at a time when measurements are already finished.

Passive cell balancing using integrated FETs is limited by low balancing current and, therefore, may require multiple cycles to correct a typical imbalance. To achieve fast passive cell balancing, an external bypass circuit can be implemented by modifying the existing hardware. Figure 4.7 shows a typical implementation. The internal balancing P-MOSFET S_N for a particular cell, which needs to be balanced, is turned on first. This creates a low-level bias current through the external resistor dividers, R1 and R2, which connect the cell terminals to the battery cell balance controller IC. The gate-to-source voltage is thus established across R2, and the external MOSFET S_{EN} is turned on. The on-resistance of the external MOSFET S_{EN} is negligible compared with the external cell balance resistance R_{BAL}, and the external balancing current, I_{BAL}, is given by $I_{BAL} = V_{CELL}/R_{BAL}$.

By properly selecting the R_{BAL} resistance value, we can get the desirable cell-balancing current, which could be much higher than the internal cell-balancing current and can speed up the cell-balancing process.

The drawback of this method is that balancing cannot be achieved on adjacent cells at the same time, as shown in Figure 4.8. When internal MOSFETs S_N and S_{N+1} of the adjacent cells are turned on, there is no net current flowing
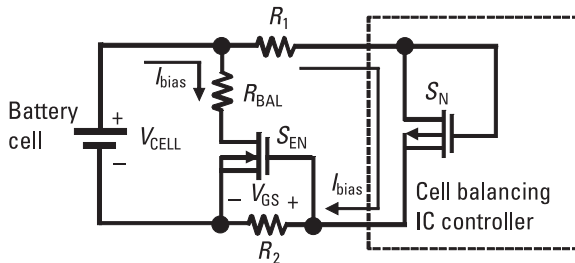

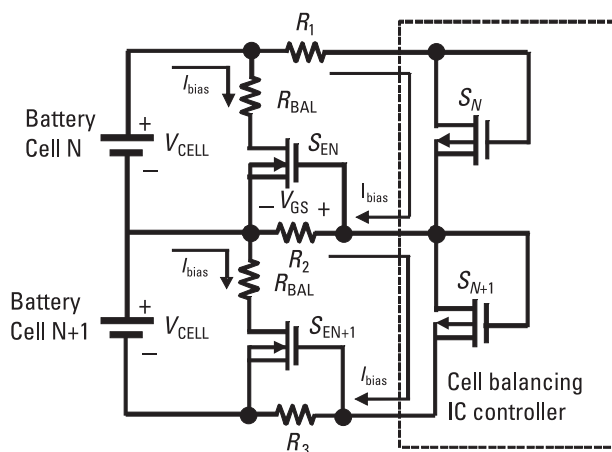**Figure 4.7**   External passive cell-balancing circuit.

Figure 4.8   The adjacent cell-balancing issue.

through R2 and no voltage drop is created across R2. So there is no gate-source bias voltage of MOSFET S_{EN}, which remains off even when the internal MOSFET S_N is on. In practice, this is not an issue because the fast external cell balancing can quickly balance the cell associated with S_{EN2}, and then the cell associated with S_{EN1} will be balanced. Therefore, the adjacent cells cannot be balanced at the same time, and only every other series cell can be balanced with this approach at the same time. However, some cell balancing ICs have two pins per cell—one for voltage sensing and one for balancing—which makes for a simpler circuit that can balance all cells simultaneously, but such ICs handle fewer cells or have a larger number of pins.

4.4.2   Charge Redistribution

The disadvantage of the current bypass approach is that the energy of the bypassed charge is wasted. While this can be acceptable during charge while the system is connected to a power grid, during actual usage of the battery in portable applications every milliwatt-hour is precious. This makes desirable a cell-balancing approach that would allow us to drain the "high" cells to the bottom in the most efficient way.

    The ultimate approach to accomplish this is to use a pack that has no serially connected cells at all. The step-up converter then ensures that the device obtains sufficient voltage. This way, energy waste as a result of the cell-balancing process is completely eliminated. The trade-off, however, is lower efficiency of the power supply, as well as increased size and complexity.

    Other solutions can include circuits that allow to for the transfer of energy from high cells to low cells rather than burning it in a bypass resistor. Note that use of the correct control scheme is still critical even in this case because

all circuits have limited efficiency and if a charge is unnecessarily shuttled back and forth multiple times due to, for example, an IR effect on voltage, overall efficiency could go close to zero after multiple "swings" even if the single-pass redistribution efficiency can be as high as 80%. So all of the balancing algorithms discussion in subsequent chapters apply to charge redistribution circuits just as they apply to bleed balancing.

4.4.3 Charge Shuttles

One simple approach for redistributing the energy among cells is to connect a capacitor first to the higher voltage cell, then to the lower voltage cell, as shown in Figure 4.9(a). More complicated implementations allow us to connect not only two nearby cells, but also several series cells, as shown in Figure 4.9(b).
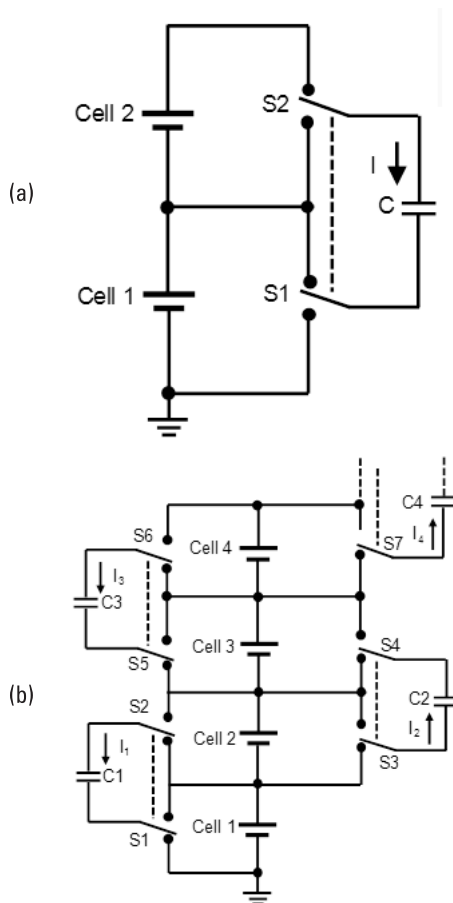

Figure 4.9 (a) Simple capacitor-based shuttle cell balancing circuit. (b) Charge shuttle circuit with several series cells.

Cell 1, cell 2, ..., cell *n* share flying capacitors with their two neighboring cells, so charge can travel from one end of the cell string to the other. This approach would take a large amount of time to transfer charge from the high cells to the low cells if they are on the opposite ends of the pack because the charge would have to travel through every cell with time and efficiency penalties. This would not be an efficient solution.

Energy loss during capacitor charging is 50%, so heating in the FETs used as switches has to be considered if high-current balancing is supported. However, because there is no charge loss with this process, the energy available on the pack terminals decreases only due to the decrease of cell voltages. Another problem is that high voltage differences between the imbalanced cells exist only in highly discharged state. Because this method transfer rate is proportional to voltage differences, it only becomes efficient near the end of discharge or the end of charge so the total amount of imbalance, that can be removed during one cycle, is low.

4.4.4   Inductive Converter–Based Cell Balancing

Active cell balancing overcomes the energy loss of the passive approach by using capacitive or inductive charge storage and shuttling energy to deliver it where it is needed most. This can be done with minimum energy loss if combined with an optimal balancing control algorithm that allows us to take full advantage of a circuit's inherent redistribution efficiency by avoiding back-and-forth shuttling. It is preferable for efficiency-conscious designs and for applications where delivering the maximum run time is top priority.

A switch-mode power converter concept can apply to the cell balance for achieving the energy transfer from one energy source to another. Figure 4.10 shows the active cell-balancing circuit based on the switch-mode power conversion concept.

A MOSFET, a diode, and a power inductor are composed of a buck-boost converter to complete a charge transfer between an adjacent pair of cells as shown in Figure 4.10. This is a bidirectional buck-boost converter, which can transfer cell energy from either direction. Figure 4.11 shows the switching waveforms of the inductor current and cell-balancing current. If the cell-balancing control algorithm determines that the top cell N needs to transfer its energy to the lower cell, the SN signal, operating at a few hundred kilohertz with a certain amount of duty cycle triggers to turn on the P-MOSFET S_N. The voltage of top cell V_{CELLN} applies to the inductor L_N and the inductor current linearly increases. The cell energy is first transferred from the top cell to the inductor during this time period. When the SN signal resets, S_N is turned off at t_1, and the energy stored in the inductor reaches a maximum value. Because the inductor current must flow continuously, the diode D_{N-1} is forward biased

Figure 4.10 Buck-boost converter based cell-balancing circuit implemented in a bq76P-L537A active cell-balancing IC.

Figure 4.11 Switching waveform of active cell-balancing circuit.

and a negative cell voltage $V_{CELLN-1}$ is applied to the inductor, which results in an inductor current decrease and transfer of the energy stored in the inductor to the lower cell. When the inductor current reaches to zero at t_2, all energy stored in the inductor has been completely transferred to the lower cell, and the diode is naturally turned off with the minimum loss. If the cell $N-1$ has more energy

than that of the top one and needs to transfer its energy to cell N, switch S_{N-1} is turned on first and the energy from cell $N-1$ is stored in inductor L. When switch S_{N-1} is turned off, then the energy stored in the inductor is transferred to the top cell N through the diode D_N. In this energy transfer process, the energy loss includes loss from the series resistance of the inductor, and the diode, and switching loss of the MOSFET.

Overall, 90% power transfer efficiency can be achieved with such active cell balancing. The balancing current is determined by the inductance switching period and its turn-on duty cycle. The current level could be much higher than passive cell balancing and more efficient. Besides the obvious advantages, the beauty of such cell-balancing technology is that balancing is achievable regardless of the individual cell voltages.

Figure 4.12 shows the active cell-balancing circuit with N series cells. From this circuit, it is found that the energy can only be transferred from the top cell to the lower adjacent cell or from the lower cell to higher cell as well.

As we know the flyback converter is the isolated power converter of the buck-boost converter. Its output is isolated from the input, and output can be floating such that it can connect anyplace. Figure 4.13 shows an active cell-balancing circuit that can transfer the energy from the bottom cell to the top cell directly with a flyback converter. When switcher S1 is turned on, the bottom cell voltage is applied to the primary winding. The current flowing through the magnetizing inductor linearly increases and its energy is stored in the magnetic field. When switcher S1 is turned off, the energy stored in the magnetizing inductor is released to charge the top cell through the output diode D_n. Therefore, the extra energy from the bottom cell can be transferred to the top cell through a flyback converter. The main limitation of such an active cell-balancing method is that energy is only distributed to the adjacent cell, not to any target cell. On the other hand, such cell-to-cell balancing is only good for battery packs with few cells in series. For long strings, due to the inefficiency of the converters at each step of the transfer, too little energy can be transferred from one end of the pack to the other, making active balancing less efficient than dissipative (passive) balancing.

How do we charge the weak cell or discharge the strong cell to achieve cell balancing? After measuring the voltage or capacity for all cells, the average voltage or capacity can be calculated. The cell with the biggest deviation from the average could be identified. The cell with the lowest voltage or capacity can be recharged from the pack, while the cell with the highest voltage or capacity could be discharged. Figure 4.14 shows the synchronous switching bidirectional flyback converter used to achieve active cell balancing for any target cell. Assume cell 2 is recognized as the weakest cell, and it needs to be charged. Switcher S_P is turned on first, the whole pack voltage is applied to the primary winding, and it stores the energy in the magnetizing inductor of the
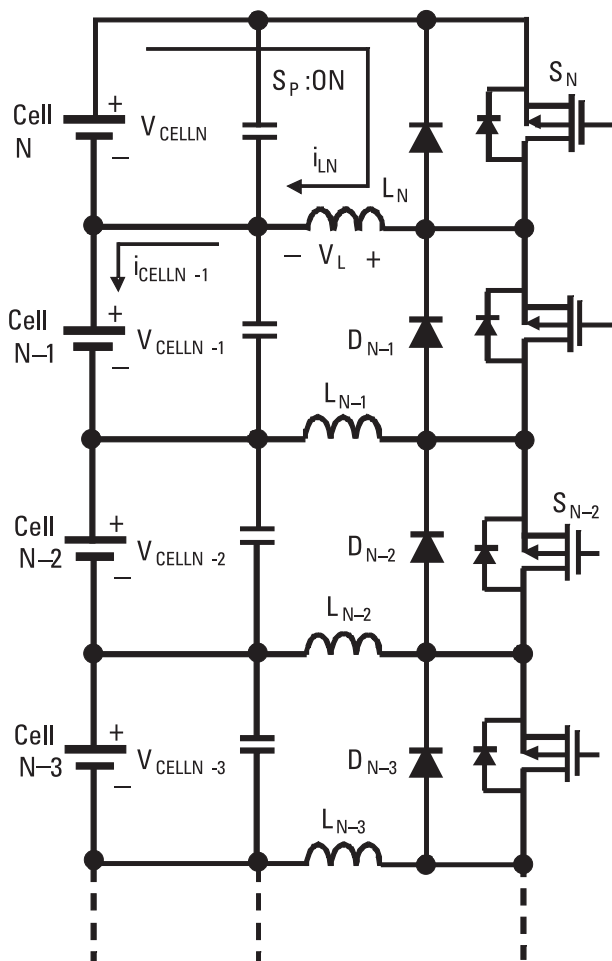
Figure 4.12   N-series cell-balancing circuit.

transformer. When switcher S_p is turned off, the MOSFET S_2 associated with the selected weakest cell is turned on. The stored energy of the transformer can be transferred to cell 2. This can extend the battery pack's run time by preventing the weakest cell from reaching the end of discharge earlier than that of the rest of cells. Therefore, this operation can achieve the cell balancing required for transferring energy from the pack to any either of the cells.

On the other hand, if is more effective to shift energy from a strong cell to the pack. Without cell balancing, the charge process has to stop immediately when one cell reaches it maximum cell voltage, even though the other cells have not been fully charged. Assume that cell $N-1$ has been detected as the strongest cell in the battery pack, and we need to discharge the energy from cell
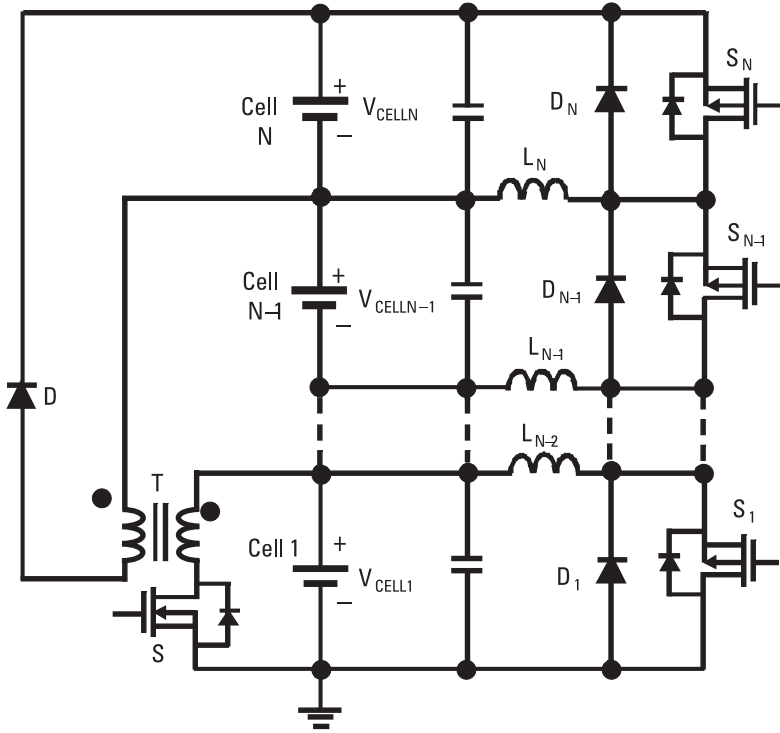
Figure 4.13 Flyback-based cell-balancing circuit for transferring the energy from the bottom cell to the top cell.

$N-1$ and redistribute it into the pack. MOSFET S_{N-1} is turned on first. The cell voltage is applied to the winding connected to cell $N-1$ and stores the energy from the cell $N-1$ in the magnetic field. Once MOSFET S_{N-1} turns off, MOSFET S_p, connected in the primary winding, is turned on so that the energy stored in the magnetic field is transferred back into the pack through the primary winding. So, this operation can achieve cell balancing by transferring cell energy to the pack.

4.5 Balancing Algorithms

Regardless of the particular hardware implementation, there is always a decision to be made regarding when to turn on a bypass switch or when to engage the energy exchange circuit for a particular cell. Different algorithms used to make this decision are reviewed in this section. For simplicity we will refer to the case of current bypass because application of the logic to other balancing schemes is trivial.
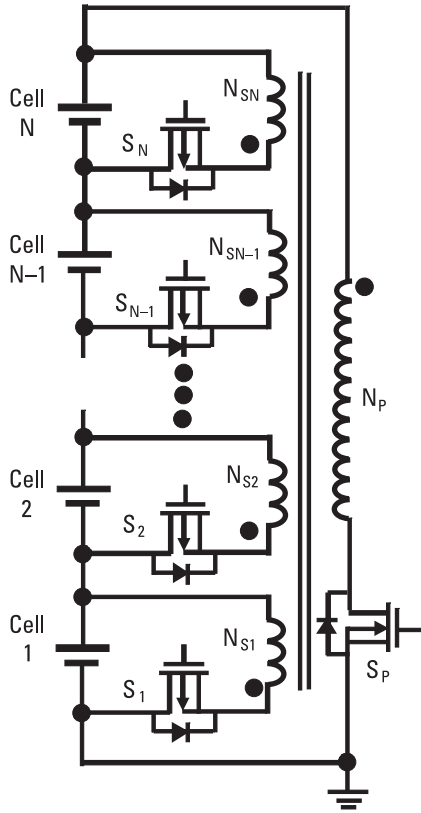
Figure 4.14 Synchronous switching bidirectional flyback converter-based active cell-balancing circuit.

4.5.1 Cell Voltage Based

The simplest algorithm is based on the voltage difference among the cells. If that difference exceeds a predefined threshold, bypass is engaged. To resolve several problems that accompany this simple method, more complicated modifications can be implemented if a microcontroller is used to execute the algorithm:

- *Balancing during charge only* is used to save energy in portable applications.

- *Balancing at high states of charge only* is used to decrease the effect on SOC balancing that can come from an impedance imbalance, because current decreases during the CV mode until charging terminates on minimal taper current.

- A variation of above, in which the balancing is enabled only at a low current regardless of the SOC.

• *Simultaneous multicell balancing* makes decisions about which cells have to be bypassed in terms of the entire pack, not just neighboring cells, as is the case with comparator-based solutions.

One advanced implementation of voltage-based algorithms that incorporates all of the above optimizations is used in the bq2084 battery fuel gauge. Figure 4.15 shows the voltage convergence of multiple cells during balancing.

4.5.2 SOC Based

If a method for determining SOC that is independent of the voltage being under load is available, the balancing algorithm can be improved, because it is no longer vulnerable to impedance variations. However, if no independent method to measure each cell's full capacity exists, equal capacity has to be assumed for all cells such that the capacity imbalance will not be considered. Such a method works as follows:

1. Determine the initial SOC for each series cell bank separately. One of the determination methods is to use an open circuit voltage correlation with the state of charge. This method can only be implemented in a microcontroller with flash memory and significant computational resources because of the need to evaluate voltage versus SOC function OCV(SOC,T) in real time.

2. Determine how much charge is needed for each cell to reach a fully charged state. This requires knowledge of total capacity, which is assumed equal for all cells.

3. Find the cell that has the largest amount of charge needed to reach full capacity, and find the differences *dQ* among all other cells that need a charge and that of the largest one.

4. This difference has to be bypassed for each "excessive" cell during one or multiple cycles. To achieve that, the bypass FETs are turned on during charging for the duration of each cell's calculated bypass time. The bypass time is calculated dependent on the value of the bypass current, which in turn depends on values of bypass resistance, R_bypass, because time = *dQ* * R_bypass/(V_average * duty_cycle). Although not a very accurate estimate of needed bypass time, this method is acceptable for low-rate bypass, because during many cycles of balancing, the amount of needed balancing time will be recalculated after every cycle.

5. Alternatively, the bypass current around each cell can be continuously integrated I = V_{cell}/R_bypass) and bypass FETs are turned off once
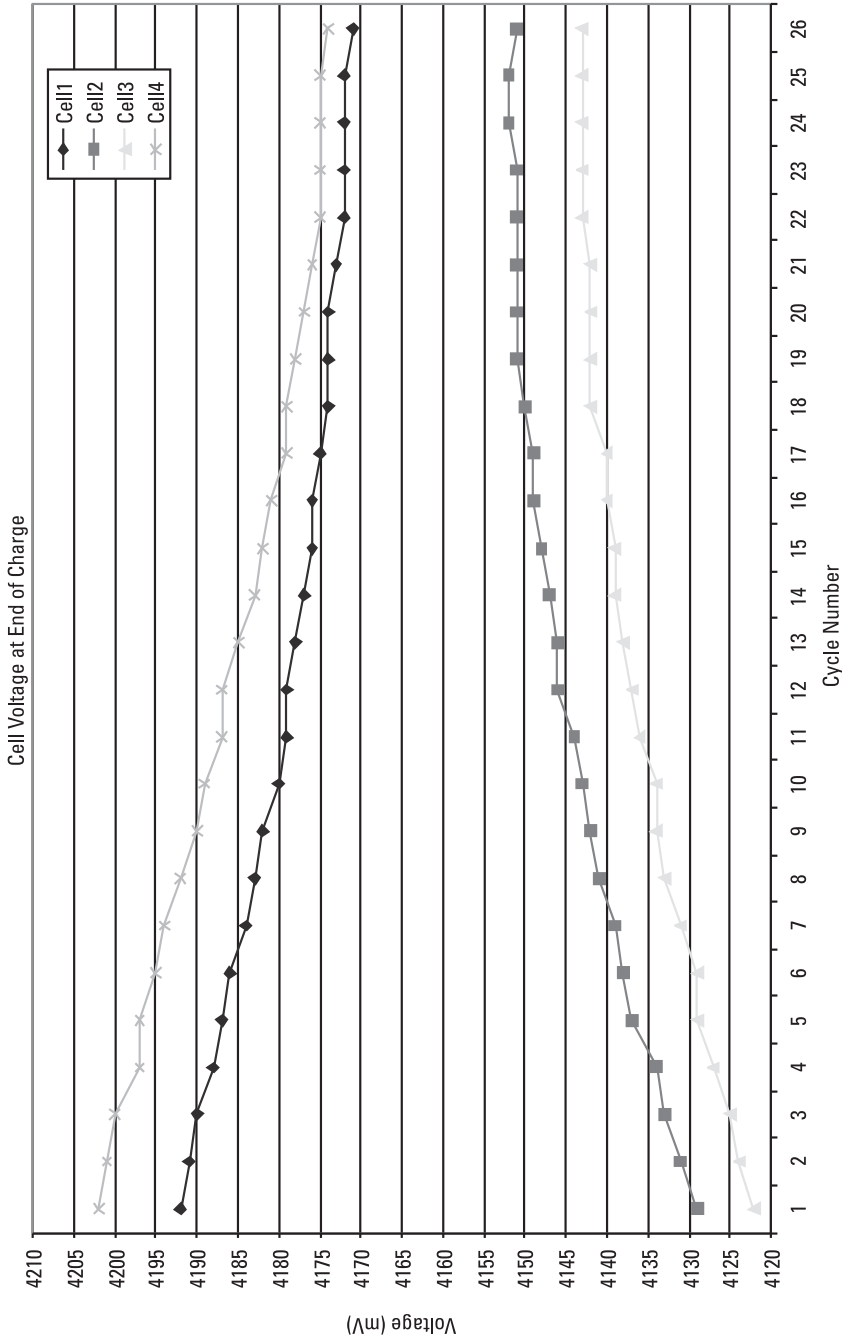
Figure 4.15   Cell open circuit voltages of a four-cell pack at the end of charge during balancing. Initial imbalance = 10%.

the needed bypass charge has passed. This method is preferable if high bypass currents, capable of balancing cells in one cycle, are used.

4.5.3   SOC and Total Capacity Based

Further improvement of the above method takes into account differences in total cell capacities (Q_{max}). This becomes possible if the cell-balancing algorithm is an integral part of the more complex gas-gauging algorithm that is monitoring the state of each cell and capable of measuring changes in the total cell capacity of each cell, as is the case, for example, with bq20z80 cell balancing. Overall, the balancing method is similar to that described in Section 4.5.2, except for calculation of dQ, which now takes into account each cell's individual capacity. Figure 4.16 shows the progress of the cell-balancing process, which causes changes in the open circuit voltage at the end of charge in a three-cell pack, where cell 1 was individually discharged and cell 2 was charged by 2% prior to the test. The bypass resistance used in this test was 700 Ω.

4.6   Summary

In this chapter, we analyzed three main mechanisms that can cause voltage differences among cells that are serially connected in a pack: (1) a SOC imbalance, (2) a total capacity imbalance, and (3) an impedance imbalance. Because the ability to add or remove only a certain amount of charge is available to
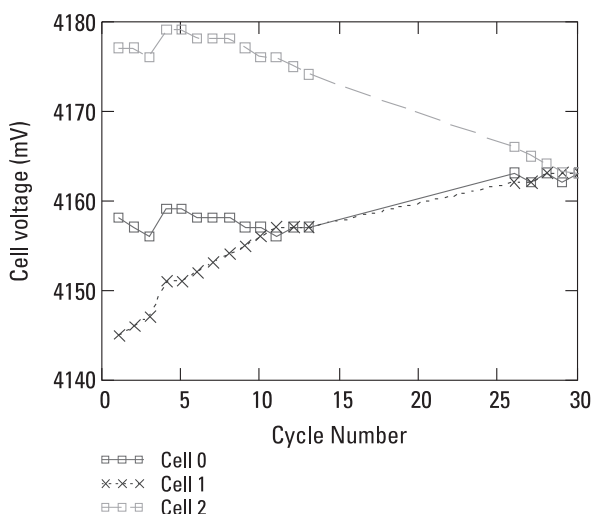

Figure 4.16   Evolution of cell voltages during SOC/Q_{max} balancing, starting from an initial 2% down (cell 1) and 2% up (cell 2) imbalance.

balancing algorithms, only the first type of imbalance, an SOC imbalance, can be eliminated. The second type of imbalance (capacity) has to be taken into account in bypass charge calculations and the third (impedance) should be kept in mind as a distortion, if voltage is used as the balancing criteria, to improve the balancing process and prevent the introduction of a larger imbalance.

Considering the low self-discharge rate of Li-ion cells, we can conclude that if continuous balancing is engaged, the use of integrated FETs provides sufficient balancing current. Use of external FETs may be required if voltage-based balancing is used; that is only active in the areas where a SOC imbalance is reflected by high voltage differences (mostly at the end of discharge). Active balancing methods can provide higher efficiency, but are not at present cost effective for portable applications. An exception might be the case in which extremely high reliability and longevity of the battery pack are needed, because active balancing extends the usable life of a pack, primarily due to the complete use of pack energy regardless of the amount of imbalance.

Voltage-based balancing algorithms have the advantage of simplicity of hardware and implementation, but suffer from slower balancing rates and the possible introduction of additional imbalance through distortions from impedance differences. SOC and total capacity based methods are more complicated to realize but can take advantage of the already present gas-gauging capability of controller ICs and ensure, for given bypass capabilities of the hardware, the most accurate and fastest balancing possible.