# 3

# Limit Order Books

*Though this be madness, yet there is method in't.*

(Shakespeare, Hamlet)

Today, most of the world's financial markets use an electronic trading mechanism called a limit order book (LOB) to facilitate trade. The Helsinki, Hong Kong, London, New York, Shenzhen, Swiss, Tokyo, Toronto and Vancouver Stock Exchanges, together with Euronext, the Australian Securities Exchange and NASDAQ, all use LOBs, as do many smaller markets. In this chapter, we provide a detailed introduction to trading via LOBs, and we discuss how price changes in an LOB emerge from the dynamic interplay between liquidity providers and liquidity takers.

## 3.1 The Mechanics of LOB Trading

In contrast to quote-driven systems (see Chapter 1), in which prices are set by designated market-makers, price formation in an LOB is a self-organised process that is driven by the submissions and cancellations of orders. Each order is a visible declaration of a market participant's desire to buy or sell a specified quantity of an asset at a specified price. Active orders reside in a **queue** until they are either cancelled by their owner or executed against an order of opposite direction. Upon execution, the owners of the relevant orders trade the agreed quantity of the asset at the agreed price.

Whereas in the old days the list of buy and sell orders was only known to the specialist, the queues of outstanding orders in LOBs are now observable in real time by traders from around the world. Because each order within these queues constitutes a firm commitment to trade, analysing the state of the LOB provides a concrete way to quantify the visible liquidity for a given asset.

### 3.1.1 Orders

An **order** is a commitment, declared at a given submission time, to buy or sell a given volume of an asset at no worse than a given price. An order $x$ is thus described by four attributes:

- its *sign* (or *direction*) $\varepsilon_x = \pm 1$, ($\varepsilon_x = +1$ for buy orders; $\varepsilon_x = -1$ for sell orders),
- its *price* $p_x$,
- its *volume* $\upsilon_x > 0$, and
- its *submission time* $t_x$.

We introduce the succinct notation

$$x := (\varepsilon_x, p_x, \upsilon_x, t_x). \tag{3.1}$$

### 3.1.2 The Trade-Matching Algorithm

Whenever a trader submits a buy (respectively, sell) order $x$, an LOB's *trade-matching algorithm* checks whether it is possible for $x$ to *match to* existing sell (respectively, buy) orders $y$ such that $p_y \leq p_x$ (respectively, $p_y \geq p_x$). If so, the matching occurs immediately and the relevant traders perform a trade for the agreed amount at the agreed price. Any portion of $x$ that does not match instead becomes *active* at the price $p_x$, and it remains active until either it matches to an incoming sell (respectively, buy) order or it is *cancelled*. **Cancellation** usually occurs when the owner of an order no longer wishes to offer a trade at the stated price.

### 3.1.3 Market Orders, Limit Orders and the LOB

Orders that match upon arrival are called **market orders**. Orders that do not match upon arrival are called **limit orders**. A **limit order book** (LOB) is simply a collection of revealed, unsatisfied intentions to buy or sell an asset at a given time. More precisely, an LOB $\mathscr{L}(t)$ is the set of all limit orders for a given asset on a given platform[1] at time $t$.

The limit orders in $\mathscr{L}(t)$ can be partitioned into the set of buy limit orders $\mathcal{B}(t)$ (for which $\varepsilon_x = +1$) and the set of sell limit orders $\mathcal{A}(t)$ (for which $\varepsilon_x = -1$).

Throughout the book, we will consider the evolution of an LOB $\mathscr{L}(t)$ as a so-called *càglàd process* (*continu à gauche, limite à droite*). Informally, this means that when a new order $x$ is submitted at time $t_x$, we regard it to be present in the LOB immediately after its arrival, but not immediately upon its arrival. More formally, we introduce the notation

$$\overline{\mathscr{L}}(t) = \lim_{t' \downarrow t} \mathscr{L}(t') \tag{3.2}$$

[1] In modern financial markets, many assets are traded on several different platforms simultaneously. The *consolidated LOB* is the union of all *LOBs* for the asset, across all platforms where it is traded.

to denote the state of $\mathscr{L}(t)$ immediately after time $t$.[2] Therefore, for a limit order $x$ submitted at time $t_x$, it holds that

$$x \notin \mathscr{L}(t_x),$$
$$x \in \overline{\mathscr{L}}(t_x).$$

Traders in an LOB are able to choose the submission price $p_x$, which will classify the order as a limit order (when it does not lead to an immediate transaction) or as a market order (when it is immediately matched to a limit order of opposite sign). Limit orders stand a chance of matching at better prices than do market orders, but they also run the risk of never being matched. Conversely, market orders match at worse prices, but they do not face the inherent uncertainty associated with limit orders. Some trading platforms allow traders to specify that they wish to submit a market order without explicitly specifying a price. Instead, such a trader specifies only a size, and the LOB's trade-matching algorithm sets the price of the order appropriately to initiate the required matching.

The popularity of LOBs is partly due to their ability to allow some traders to demand immediacy (by submitting market orders), while simultaneously allowing others to supply it (by submitting limit orders) and hence to at least partially play the traditional role of market-makers. Importantly, liquidity provision in an LOB is a decentralised and self-organised process, driven by the submission of limit orders from all market participants. In reality, many traders or execution algorithms use a combination of both limit orders and market orders by selecting their actions for each situation based on their individual needs at that time.

The following remark can be helpful in thinking about the relative merits of limit orders and market orders. In an LOB, each limit order can be construed as an option contract written to the whole market, via which the order's owner offers to buy or sell the specified quantity $\upsilon_x$ of the asset at the specified price $p_x$ to any trader wishing to accept. For example, a trader who submits a sell limit order $x = (-1, p_x, \upsilon_x, t_x)$ is offering the entire market a free option to buy $\upsilon_x$ units of the asset at price $p_x$ for as long as the order remains active. Traders offer such options – i.e. submit limit orders – in the hope that they will be able to trade at better prices than if they simply submitted market orders. However, whether or not a limit order will eventually become matched is uncertain, and if it is matched, there is a good chance that the price will continue drifting beyond the limit price, leading to a loss for the issuer of the option. This is an example of adverse selection and its associated skewness, which we already discussed in Section 1.3.2. In a nutshell, as we will show later in the book (see Section 21.3), limit orders typically earn profits when prices mean-revert but suffer losses when prices trend.

[2] The notation $t' \downarrow t$ ($t' \uparrow t$) means that $t'$ approaches $t$ from above (below).

### 3.1.4 The Bid-, Ask- and Mid-Price

The terms *bid-price*, *ask-price*, *mid-price* and *spread* (all of which we have already encountered in Chapter 1) are common to much of the finance literature. Their definitions can be made specific in the context of an LOB (see Figure 3.1):

- The **bid-price** at time $t$ is the highest stated price among buy limit orders at time $t$,

$$b(t) := \max_{x \in \mathcal{B}(t)} p_x. \tag{3.3}$$

- The **ask-price** at time $t$ is the lowest stated price among sell limit orders at time $t$,

$$a(t) := \min_{x \in \mathcal{A}(t)} p_x. \tag{3.4}$$

At any given time, $b(t)$ is therefore the highest price at which it is immediately possible to sell and $a(t)$ is the lowest price at which it is immediately possible to buy. By definition, $a(t) > b(t)$, otherwise some buy and sell limit orders can be immediately matched and removed from the LOB.

- The **mid-price** at time $t$ is

$$m(t) := \frac{1}{2}\left[a(t) + b(t)\right].$$

- The **bid–ask spread** (or simply "spread") at time $t$ is

$$s(t) := a(t) - b(t) > 0.$$

Each of $b(t)$, $a(t)$, $m(t)$ and $s(t)$ are also càglàd processes. We use the same overline notation as in Equation (3.2) to denote the values of these processes immediately after time $t$. For example,

$$\overline{m}(t) = \lim_{t' \downarrow t} m(t')$$

denotes the mid-price immediately after $t$.

Figure 3.1 shows a schematic of an LOB at some instant in time, illustrating the above definitions. The horizontal lines within the blocks at each price level denote how the total volume available at that price is composed of different limit orders. As the figure illustrates, an LOB can be regarded as a set of queues, each of which consists of a set of buy or sell limit orders at a specified price. As we will discuss at many points throughout the book, this interpretation of an LOB as a system of queues often provides a useful starting point for building LOB models.

### 3.1.5 The Lot Size and Tick Size

When submitting an order $x$, a trader must choose the size $\upsilon_x$ and price $p_x$ according to the relevant lot size and tick size of the given LOB.

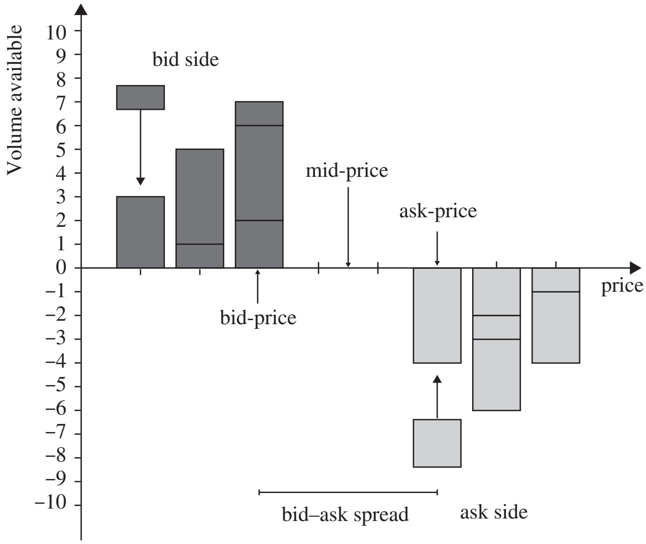


Figure 3.1. A schematic of an LOB, to illustrate the bid-price, the ask-price, the available volumes, the mid-price, and the bid–ask spread.

- The **lot size** $\upsilon_0$ is the smallest amount of the asset that can be traded within the given LOB. All orders must arrive with a size

$$\upsilon_x \in \{k\upsilon_0 \,|k = 1, 2, \ldots\}.$$

- The **tick size** $\vartheta$ is the smallest permissible price interval between different orders within a given LOB. All orders must arrive with a price that is a multiple of $\vartheta$.

The lot size and tick size fix the units of order size and order price in a given LOB. For example, if $\vartheta$ = \$0.001, then the largest permissible order price that is strictly less than \$1.00 is \$0.999, and all orders must be submitted at a price with exactly three decimal places. The lot size $\upsilon_0$ and tick size $\vartheta$ of an LOB are collectively called its *resolution parameters.*

Values of $\upsilon_0$ and $\vartheta$ vary greatly from between different trading platforms. Expensive stocks are often traded with $\upsilon_0 = 1$ share; cheaper stocks are often traded with $\upsilon_0 \gg 1$ share. In foreign exchange (FX) markets, some trading platforms use values as large as $\upsilon_0 = 1$ million units of the base currency, whereas others use values as small as $\upsilon_0 = 0.01$ units of the base currency.[3] In equity markets, $\vartheta$ is often 0.01% of the stock's mid-price $m(t)$, rounded to the nearest power of 10. In US equity markets, the tick size is always $\vartheta$ = \$0.01, independently of the stock price. A given asset is sometimes traded with different values of $\vartheta$ on different

[3] In FX markets, an XXX/YYY LOB matches exchanges of the *base currency* XXX to the *counter currency* YYY. A price in an XXX/YYY LOB denotes how many units of currency YYY are exchanged for a single unit of currency XXX. For example, a trade at the price \$1.28124 in a GBP/USD market corresponds to 1 pound sterling being exchanged for 1.28124 US dollars.

trading platforms. For example, on the electronic trading platform Hotspot FX, $\vartheta$ = \$0.00001 for the GBP/USD LOB and $\vartheta$ = 0.001 for the USD/JPY LOB, whereas on the electronic trading platform EBS, $\vartheta$ = \$0.00005 for the GBP/USD LOB and $\vartheta$ = 0.005 for the USD/JPY LOB.

An LOB's resolution parameters greatly affect trading. The lot size $\upsilon_0$ dictates the smallest permissible order size, so any trader who wishes to trade in quantities smaller than $\upsilon_0$ is unable to do so. Furthermore, as we will discuss in Chapter 10, traders who wish to submit large market orders often break them into smaller chunks to minimise their price impact. The size of $\upsilon_0$ controls the smallest permissible size of these chunks and therefore directly affects traders who implement such a strategy.

The tick size $\vartheta$ dictates how much more expensive it is for a trader to gain the priority (see Section 3.2.1) associated with choosing a higher (respectively, lower) price for a buy (respectively, sell) order. In markets where $\vartheta$ is extremely small, there is little reason for a trader to submit a buy (respectively, sell) limit order at a price $p$ where there are already other limit orders. Instead, the trader can gain priority over such limit orders very cheaply, by choosing the price $p + \vartheta$ (respectively, $p - \vartheta$) for their limit order. Such a setup leads to very small volumes at any level in the LOB, including the best quotes $b(t)$ and $a(t)$, and therefore leads to extremely frequent changes in $b(t)$ and $a(t)$.

Some market commentators argue that small tick sizes make it difficult for traders to monitor the state of the market in real time. In September 2012, the electronic FX trading platform EBS increased the size of $\vartheta$ for most of its currency pairs, to "help thicken top of book price points, increase the cost of top-of-book price discovery, and improve matching execution in terms of percent fill amounts". However, as we will see at many times throughout the book, an asset's tick size can influence order flow in many different ways, some of which are quite surprising. Therefore, understanding how changing the tick size will influence future market activity is far from straightforward.

Even for LOBs with the same or similar resolution parameters $\upsilon_0$ or $\vartheta$, these can represent vastly different fractions of the typical trade size and price. For example, both the Priceline Group and the Amyris Inc. stocks are traded on NASDAQ with $\vartheta$ = \$0.01, yet the typical price for Priceline exceeds \$1000.00 whereas the typical price for Amyris is close to \$1.00. Therefore, $\vartheta$ constitutes a much larger fraction of the typical trade size for Amyris than it does for Priceline. For this reason, it is sometimes useful to consider the **relative tick size** $\vartheta_r$, which is equal to the tick size $\vartheta$ divided by the mid-price $m(t)$ for the given asset. For example, the price of Priceline Group was on the order of \$1000 in 2014, corresponding to a very small relative tick size $\vartheta_r$ of $10^{-5}$.

### 3.1.6 Relative Prices

Because the activity of a single market participant is driven by his/her trading needs, individual actions can appear quite erratic. However, when measured in a suitable coordinate frame that aggregates order flows from many different market participants, robust statistical properties of order flow and LOB state can emerge from the ensemble – just as the ideal gas law emerges from the complicated motion of individual molecules.

Most studies of LOBs perform this aggregation in **same-side quote-relative coordinates**, in which prices are measured relative to the same-side best quote. Specifically, the same-side quote-relative price of an order $x$ at time $t$ is (see Figure 3.1)

$$d(p_x,t) := \begin{cases} b(t) - p_x, & \text{if } x \text{ is a buy limit order,} \\ p_x - a(t), & \text{if } x \text{ is a sell limit order.} \end{cases} \tag{3.5}$$

In some cases (such as when measuring volume profiles, as in Section 4.7), using same-side quote-relative coordinates can cause unwanted artefacts to appear in statistical results. Therefore, instead of measuring prices relative to the same-side best quote, it is sometimes more useful to measure prices in **opposite-side quote-relative coordinates**. The opposite-side quote-relative price of an order $x$ at time $t$ is (see Figure 3.1)

$$d^{\dagger}(p_x,t) := \begin{cases} a(t) - p_x, & \text{if } x \text{ is a buy limit order,} \\ p_x - b(t), & \text{if } x \text{ is a sell limit order.} \end{cases} \tag{3.6}$$

The difference in signs between the definitions for buy and sell orders in Equations (3.5) and (3.6) ensures that an increasing distance from the reference price is always recorded as positive. By definition, all limit orders have a non-negative same-side quote-relative price and a strictly positive opposite-side quote-relative price at all times. Same-side and opposite-side quote-relative prices are related by the bid–ask spread $s(t)$:

$$d^{\dagger}(p_x,t) = d(p_x,t) + s(t).$$

The widespread use of quote-relative coordinates is motivated by the notion that market participants monitor $b(t)$ and $a(t)$ when deciding how to act. There are many reasons why this is the case. For example, $b(t)$ and $a(t)$ are observable to all market participants in real time, are common to all market participants, and define the boundary conditions that dictate whether an incoming order $x$ is a limit order ($p_x < a(t)$ for a buy order or $p_x > b(t)$ for a sell order) or a market order ($p_x \geq a(t)$ for a buy order or $p_x \leq b(t)$ for a sell order). Therefore, the bid and the ask constitute suitable reference points for aggregating order flows across different market participants.

Quote-relative coordinates also provide a useful benchmark for understanding LOB activity. When studying LOBs, it is rarely illuminating to consider the actual price $p_x$ of an order, because this information provides no context for $x$ in relation to the wider activity in $\mathscr{L}(t)$. As we show in Chapter 4, by instead studying the quote-relative price of an order, it is possible to understand the role of the order in relation to the other orders in the market.

### 3.1.7 The Volume Profile

Most traders assess the state of $\mathscr{L}(t)$ via the **volume profile** (or "depth profile"). The buy-side volume at price $p$ and at time $t$ is

$$V_+(p,t) := \sum_{\{x \in \mathscr{B}(t) | p_x = p\}} \upsilon_x. \tag{3.7}$$

The sell-side volume at price $p$ and at time $t$, denoted $V_-(p,t)$, is defined similarly using $\mathscr{A}(t)$. The volume profile at time $t$ is the set of volumes at all prices $p$, $\{V_\pm(p,t)\}$.

Because $b(t)$ and $a(t)$ vary over time, quote-relative coordinates provide a useful approach to studying the volume profile through time, akin to changing reference frame in physics.

### 3.1.8 Price Changes in an LOB

In an LOB, the rules that govern matching dictate how prices evolve through time. Consider a buy (respectively, sell) order $x = (\pm 1, p_x, \upsilon_x, t_x)$ that arrives at time $t$.

- If $p_x \leq b(t)$ (respectively, $p_x \geq a(t)$), then $x$ is a limit order. It does not cause $b(t)$ or $a(t)$ to change.
- If $b(t) < p_x < a(t)$, then $x$ is also a limit order. It causes $b(t)$ to increase (respectively, $a(t)$ to decrease) to $p_x$ at time $t_x$.
- If $p_x \geq a(t)$ (respectively, $p_x \leq b(t)$), then $x$ is a market order that matches to one or more buy (respectively, sell) limit orders upon arrival. Whenever such a matching occurs, it does so at the price of the limit orders, which can lead to "price improvement" (compared to the price $p_x$ of the incoming market order). Whether or not such a matching causes $b(t)$ (respectively, $a(t)$) to change at time $t_x$ depends on the volume available at the bid $V_+(b(t_x),t)$ (respectively, at the ask $V_-(a(t_x),t)$) compared to $\upsilon_x$. In particular, the new bid-price $\overline{b}(t_x)$ immediately after the arrival of a sell market order $x$ is set as follows. One first computes the *largest* price $b^*$ such that such that all the volume between $b^*$ and $b(t_x)$ is greater

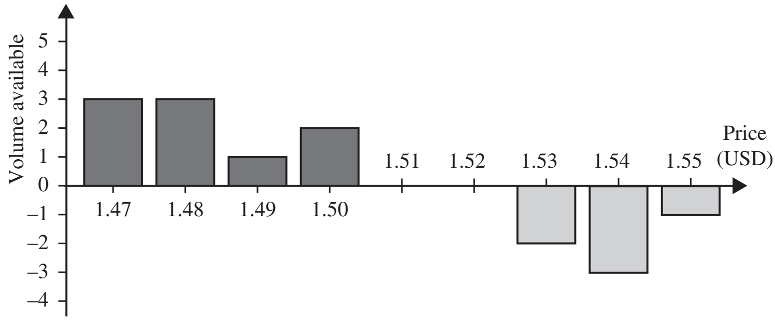


Figure 3.2. A specific example of an LOB, as presented in Table 3.1.

than or equal to $\upsilon_x$:

$$\sum_{p=b^*}^{b(t_x)} V_+(p,t) \geq \upsilon_x,$$

$$\sum_{p=b^*+\vartheta}^{b(t_x)} V_+(p,t) < \upsilon_x.$$

Then $\overline{b}(t_x)$ is equal to $b^*$ whenever $p_x \leq b^*$ and to the limit price $p_x$ otherwise. In the latter case, the market order is not fully executed.

Similarly, the new ask-price $\overline{a}(t_x)$ immediately after the arrival of a buy market order $x$ is the smallest of $a^*$ or $p_x$, where:

$$\sum_{p=a(t_x)}^{a^*} V_-(p,t) \geq \upsilon_x,$$

$$\sum_{p=a(t_x)}^{a^*-\vartheta} V_-(p,t) < \upsilon_x.$$

To illustrate the above expressions, Table 3.1 lists several possible market events that could occur to the LOB displayed in Figure 3.2 and the resulting values of $\overline{b}(t_x), \overline{a}(t_x), \overline{m}(t_x)$ and $\overline{s}(t_x)$ after their arrival. Figure 3.3 shows how the arrival of an order $(+1, \$1.55, 3, t_x)$, as described by the third line of Table 3.1, would impact the LOB shown in Figure 3.2.

## 3.2 Practical Considerations

Many practical details of trading vary considerably across LOB platforms. In this section, we highlight some of these practical differences and discuss how they can influence both the temporal evolution of an LOB and the actions of traders within it.

Table 3.1. *Example to illustrate how specified order arrivals would affect prices in the LOB in the top panel of Figure 3.3. The tick size is* $\vartheta = 0.01$ *and the lot size is* $\upsilon_0 = 1$ *(see Section 3.1.5).*

| Arriving order $x$ | Values before arrival (USD) | | | | Values after arrival (USD) | | | |
|---|---|---|---|---|---|---|---|---|
| $(\varepsilon_x, p_x, \upsilon_x, t_x)$ | $b(t_x)$ | $a(t_x)$ | $m(t_x)$ | $s(t_x)$ | $\overline{b}(t_x)$ | $\overline{a}(t_x)$ | $\overline{m}(t_x)$ | $\overline{s}(t_x)$ |
| $(+1, \$1.48, 3, t_x)$ | 1.50 | 1.53 | 1.515 | 0.03 | 1.50 | 1.53 | 1.515 | 0.03 |
| $(+1, \$1.51, 3, t_x)$ | 1.50 | 1.53 | 1.515 | 0.03 | 1.51 | 1.53 | 1.52 | 0.02 |
| $(+1, \$1.55, 3, t_x)$ | 1.50 | 1.53 | 1.515 | 0.03 | 1.50 | 1.54 | 1.52 | 0.04 |
| $(+1, \$1.55, 5, t_x)$ | 1.50 | 1.53 | 1.515 | 0.03 | 1.50 | 1.55 | 1.525 | 0.05 |
| $(-1, \$1.54, 4, t_x)$ | 1.50 | 1.53 | 1.515 | 0.03 | 1.50 | 1.53 | 1.515 | 0.03 |
| $(-1, \$1.52, 4, t_x)$ | 1.50 | 1.53 | 1.515 | 0.03 | 1.50 | 1.52 | 1.51 | 0.02 |
| $(-1, \$1.47, 4, t_x)$ | 1.50 | 1.53 | 1.515 | 0.03 | 1.48 | 1.53 | 1.505 | 0.05 |
| $(-1, \$1.50, 4, t_x)$ | 1.50 | 1.53 | 1.515 | 0.03 | 1.49 | 1.50 | 1.495 | 0.01 |

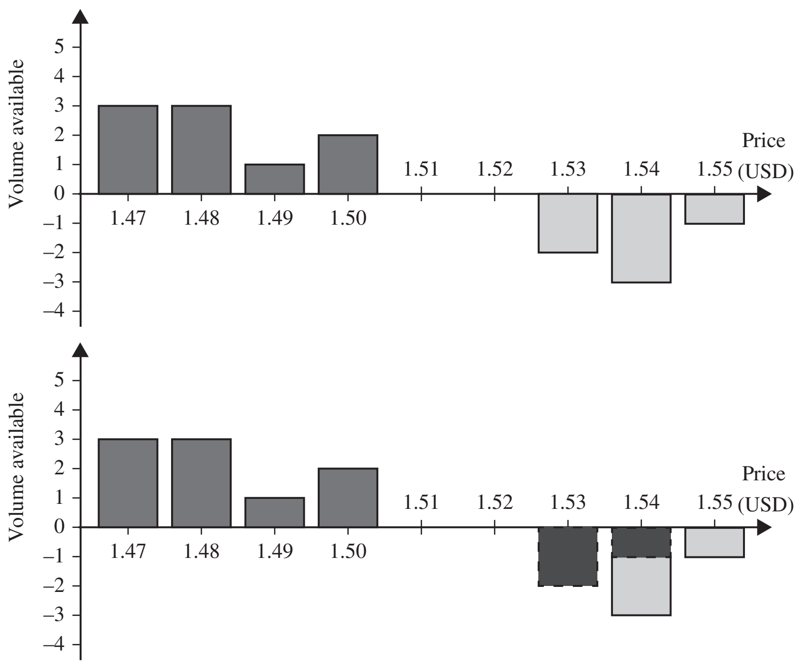


Figure 3.3. An illustration of how the arrival of an order $(+1, \$1.55, 3, t_x)$, as described by the third line of Table 3.1, would impact the LOB shown in Figure 3.2. The dashed boxes denote the limit orders that match to the incoming market order, and that are therefore removed from the LOB. The ask-price immediately after the market order arrival is $\overline{a}(t_x) = \$1.54$.

### 3.2.1 Priority Rules

As shown in Figure 3.1, several different limit orders can reside at the same price at the same time. Much like priority is given to limit orders with the best (i.e. highest buy or lowest sell) price, LOBs also employ a *priority system* for limit orders within each individual price level.

By far the most common rule currently used is **price-time priority**. That is, for buy (respectively, sell) limit orders, priority is given to the limit orders with the highest (respectively, lowest) price, and ties are broken by selecting the limit order with the earliest submission time $t_x$.

Another priority mechanism, commonly used in short-rate futures markets, is **pro-rata priority**. Under this mechanism, when a tie occurs at a given price, each relevant limit order receives a share of the matching equal to its fraction of the available volume at that price. Traders in pro-rata priority LOBs are faced with the substantial difficulty of optimally selecting limit order sizes: posting limit orders with larger sizes than the quantity that is really desired for trade becomes a viable strategy to gain priority.

Different priority mechanisms encourage traders to behave in different ways. Price-time priority encourages traders to submit limit orders early (even at prices away from the best quotes) and increase the available liquidity. Indeed, without a priority mechanism based on time, there is no incentive for traders to show their hand by submitting limit orders earlier than is absolutely necessary. Pro-rata priority rewards traders for placing larger limit orders and thus for providing greater liquidity at the best quotes. Because they directly influence the ways in which traders act, priority mechanisms play an important role, both for models and for market regulation.

### 3.2.2 Order Types

The actions of traders in an LOB can be expressed solely in terms of the submission or cancellation of orders of elementary size $\upsilon_0$. For example, a trader who sends a sell market order of $4\upsilon_0$ units of the traded asset in the LOB displayed in Figure 3.2 can be regarded as submitting two sell orders of size $\upsilon_0$ at the price \$1.50, one sell order of size $\upsilon_0$ at the price \$1.49, and one sell order of size $\upsilon_0$ at the price \$1.48. Similarly, a trader who posts a sell limit order of size $4\upsilon_0$ at the price \$1.55 can be regarded as submitting four sell orders of size $\upsilon_0$ at a price of \$1.55 each.

Because all traders' actions can be decomposed in this way, it is customary to study LOBs in terms of these simple building blocks. In practice, many platforms offer traders the ability to submit a wide assortment of order types, each with complicated rules regarding its behaviour. Although such orders are rarely discussed in the literature on LOBs (because it is possible to decompose the resulting order flow into elementary limit and/or market orders), in practice their use is relatively widespread. We therefore provide a brief description of the main order types.[4]

- **Stop orders**: A stop order is a buy (respectively sell) market order that is sent as soon as the price of a stock reaches a specified level above (respectively below) the current price. This type of order is used by investors who want to avoid big losses or protect profits without having to monitor the stock performance. Stop orders can also be limit orders.
- **Iceberg orders**: Investors who wish to submit a large limit order without being detected can submit an iceberg order. For this order type, only a fraction of the order size, called *peak volume*, is publicly disclosed. The remaining part is not visible to other traders, but usually has lower time priority, i.e. regular limit orders at the same price that are submitted later will be executed first. When the publicly disclosed volume is filled and a hidden volume is still available, a new peak volume enters the book. On exchanges where iceberg orders are allowed, their use can be quite frequent. For example, according to studies of iceberg orders on Euronext, 30% of the book volume is hidden.
- **Immediate-or-cancel orders**: If a market participant wants to profit from a trading opportunity that s/he expects will last only a short period of time, s/he can send an "immediate-or-cancel" (IOC) order. Any volume that is not matched is cancelled, such that it never enters the LOB, and leaves no visible trace unless it is executed. In particular, if zero volume is matched, everything is as if the IOC order had never been sent. This order type is similar to the all-or-nothing (or fill-or-kill) order. Orders of the latter are either executed completely or are not executed at all. In this way, the investor avoids revealing his/her intention to trade if the entire quantity was not available.
- **Market-on-close orders**: A market-on-close (MOC) order is a market order that is submitted to execute as close to the closing price as possible. If a double-sided closing auction takes place at the end of the day (such as on the New York Stock Exchange (NYSE) or the Tokyo Stock Exchange), the order will participate in the auction. Investors might want to trade at the close because they expect liquidity to be high.

### *3.2.3 Opening and Closing Auctions*

Many exchanges are closed overnight and suspend standard LOB trading at the beginning and end of the trading day. During these two periods, volumes are so large that LOB trading would be prone to instabilities, so exchanges prefer to use an auction system to match orders. For example, the LSE's flagship order book SETS has three distinct trading phases in each trading day:

[4] From Kockelkoren, J. (2010). *Encyclopedia of quantitative finance*. Wiley.

- a ten-minute **opening auction** between 07:50 and 08:00;
- a **continuous trading** period between 08:00 and 16:30 (during which the standard LOB mechanism is used); and
- a five-minute **closing auction** between 16:30 and 16:35.

During the opening and closing auctions, traders can place orders as usual, but no orders are matched. Due to the absence of matching, the highest price among buy orders can exceed the lowest price among sell orders. All orders are stored until the auction ends. At this time, for each price $p$ at which there is non-zero volume available, the trade-matching algorithm calculates the total volume $Q(p)$ of trades that could occur by matching buy orders with a price greater than or equal to $p$ to sell orders with a price less than or equal to $p$ (this is precisely the mechanism described by Equation (1.4)). It then calculates the **auction price**:

$$p^* = \arg\max_p Q(p). \tag{3.8}$$

In contrast to standard LOB trading, all trades take place at the same price $p^*$. Given $p^*$, if there is a smaller volume available for sale than there is for purchase (or vice-versa), ties are broken using time priority.

Throughout the opening and closing auctions, all traders can see what the value of $p^*$ would be if the auction were to end at that moment. This allows all traders to observe the evolution of the price without any matching taking place until the process is complete, and to revise their orders if needed. Such a price-monitoring process is common to many markets.

**Take-Home Messages**

(i) Most modern markets use limit order books (LOBs) to facilitate trade. LOBs allow liquidity takers to conduct transactions with the liquidity posted by liquidity providers.

(ii) In an LOB, traders interact via orders. An order consists of a direction (buy/sell), a price, a volume, and a submission time. The price and volume must be multiples of the tick and lot sizes, respectively.

(iii) Patient orders (i.e. orders that do not trigger an immediate transaction, and that are therefore added to the LOB) are called limit orders. Impatient orders (i.e. orders that trigger an immediate transaction against an existing limit order) are called market orders.

(iv) At any given time, the state of an LOB is simply the set of all the limit orders. The price of the best buy limit order is called the bid-price $b$, and the price of the best sell limit order is called the ask-price $a$.

(v) The quantity $m := (b + a)/2$ is called the mid-price. The quantity $s := a - b$ is called the bid–ask spread.

(vi) The mid-price and bid–ask spread both change whenever the bid-price or ask-price change. Such changes can be caused by the disappearance of the corresponding limit orders, or by the arrival of a new limit order inside the spread.

(vii) When a market order arrives, limit orders are executed according to priority rules. Limit orders with better prices (i.e. a higher buy price or lower sell price) always have priority over limit orders with worse prices. In case of price equality, other rules (such as time- or size-priority) are used to break the tie. In this way, an LOB can be regarded as a queuing system, with limit orders residing in queues at different prices.

## 3.3 Further Reading


Stoll, H. R. (1978). The pricing of security dealer services: An empirical study of NASDAQ stocks. *The Journal of Finance*, 33(4), 1153–1172.

Cohen, K. J., Conroy, R. M., & Maier, S. F. (1985). Order flow and the quality of the market. In *Market-making and the changing structure of the securities industry* (pp. 93–110). Lexington Books.

Glosten, L. R. (1994). Is the electronic open limit order book inevitable? *The Journal of Finance*, 49(4), 1127–1161.

Handa, P., & Schwartz, R. A. (1996). Limit order trading. *The Journal of Finance*, 51(5), 1835–1861.

Harris, L. (2003). *Trading and exchanges: Market microstructure for practitioners*. Oxford University Press.

Biais, B., Glosten, L., & Spatt, C. (2005). Market microstructure: A survey of microfoundations, empirical results, and policy implications. *Journal of Financial Markets*, 8(2), 217–264.

Hasbrouck, J. (2007). *Empirical market microstructure: The institutions, economics, and econometrics of securities trading*. Oxford University Press.

Parlour, C. A., & Seppi, D. J. (2008). Limit order markets: A survey. In Thakor, A. V. and Boot, A. (Eds.), *Handbook of financial intermediation and banking* (Vol. 5, pp. 63–95). North-Holland.

Foucault, T., Pagano, M., & Röell, A. (2013). *Market liquidity: Theory, evidence, and policy*. Oxford University Press.

Lehalle, C. A., & Laruelle, S. (2013). *Market microstructure in practice*. World Scientific.