.

A General Internal Fee Structure (GIFS) For a Broad Class Of AMMs.

(Or *A Glorious Internal Fee Structure*)

In AMM liquidity pools, fees may be stored *internally* in the pool or *externally*. The former is standard for Constant Product pools, while the latter is standard for Concentrated Liquidity pools, and there are advantages and disadvantages to each. While external fees are fairly trivial to implement in most contexts, using internal fees is not as easy. For example, no method is currently available to use internal fees for concentrated liquidity pools. In this document, we will design a mathematically consistent method for implementing internal fees in a relatively broad class of possible AMMs, and in a follow-up paper, we apply the method directly to the concentrated liquidity AMM. This will be divided into 3 sections:

- Section 1: Motivation & Main Idea
- Section 2: Deriving the GIFS
- Section 3: Algorithm Summary & Discussion

The mathematical derivation is quite involved. The reader who desires only the main idea and results is encouraged to read the first and last page, skipping all of Section 2. Otherwise, enjoy.

Section 1: Motivation & Main Idea

The purpose of this document is to introduce a novel mechanism for taking fees in a fairly general class of AMMs. In this opening section, we will first review existing fee structures, discuss some of the advantages and disadvantages, and outline the main idea of our novel fee structure, before diving into the mathematics in Section 2.

First, let us articulate the difference between *internal* and *external* fees. This topic is discussed in more detail in this paper, but we will quickly review it here for completeness. We consider a constant product pool between two types of assets x and y, the amounts of which are constrained to exist on the curve $xy = L^2$ in the x/y-plane, and we consider an amount Δx to be paid into the pool. A fraction γ (the *fee parameter*) splits this quantity into two pieces, $\gamma\Delta x$ and (1-γ)Δx, the latter being the fee (Figure 1).

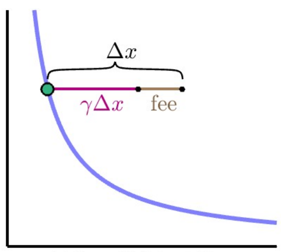


Figure 1. An incoming quantity Δx.

The quantity $\gamma\Delta x$ is used to determine the corresponding output quantity Δy via the constant product protocol, illustrated in Figure 2 (left). In an *external* fee structure, the remaining fee would then be delivered to the LP. In an *internal* fee structure, on the other hand, the entire input quantity Δx (including the fee portion) is deposited into the pool (Figure 2, right). Because the output quantity Δy was computed using a modified input quantity $\gamma\Delta x$, we are breaking the constant product rule. Rather, we graduate to a *new* curve corresponding to a slightly higher value for the liquidity constant L.

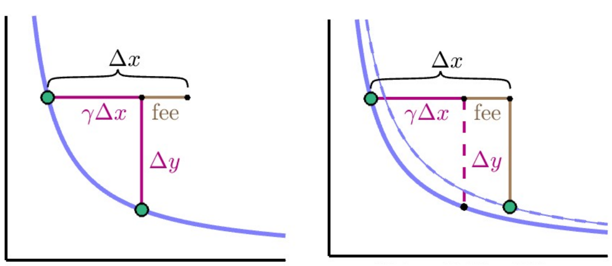


Figure 2. External fee (left) and internal fee (right).

There are advantages and disadvantages for each of these methods. Storing fees internally allows the fees to compound, and under certain market conditions this effect is non-trivial. On the other hand, fees stored externally are not exposed to impermanent loss. There may also be differences in the gas costs. See the paper cited above for a deeper discussion. We note that for constant product pools, *internal* fees are the standard.

In the context of *concentrated liquidity,* we have a sequence of active price ticks $\{p_i\}$ that establish price subranges $[p_i, p_{i+1}]$, each one of which having its own liquidity scale L_i. It is standard to use *external* fees in this case, where a fee is removed from the incoming quantity and awarded to the relevant LPs (pro rata). This requires us to store fee variables that scale with the number of active ticks. One might want to eliminate these extra fee variables by storing the fees internally in the pool. However, if we naively import that internal fee method described prior (for a constant product pool), this would not agree with the structure of the LP positions for concentrated liquidity.

Specifically, the liquidity scale factor L_i corresponding to the range $[p_i, p_{i+1}]$ is the *sum* of all LP liquidities that overlap with this range. For a trade that moves the price by Δp, there is a strict algebra relating $(L_i, \Delta p)$ to the incoming/outgoing amounts $(\Delta x, \Delta y)$. If we take a fee from the incoming quantity and deposit it back into the pool, this algebra would be broken, and a complicated rewriting of all the chosen LP positions would be necessary to maintain mathematical cohesion. A more detailed discussion of this problem is found here.

Instead, we present an alternative approach to storing the fee internally. First, we calculate the appropriate $(\Delta x, \Delta y)$ as if *no fees were taken.* Then, we simply scale up the liquidity L by some factor η. We will refer to this approach as a *fee-by-scaling* method. This is illustrated in Figure 3, on a curve representing some range $[p_i, p_{i+1}]$.

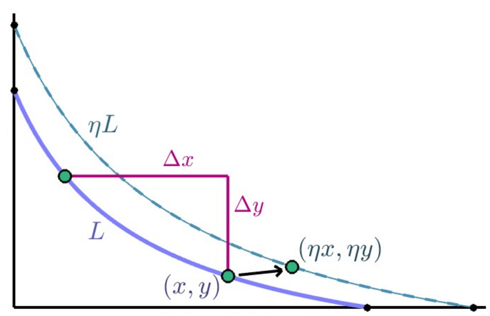


Figure 3. The *fee-by-scaling* method.

Scaling the liquidity parameter L in this way results in a commensurate scaling of the coordinates (x, y), but leaves the final price p unchanged. One can check that this is *not* the case for the internal fee structure illustrated in Figure 2. While it may not be immediately obvious, *fee-by-scaling* is enormously useful for managing the mathematical nuances of concentrated liquidity.

The obvious next question then is the choice of the scaling factor η. We will spend a great deal of time deriving our formulas, but the end result is surprisingly simple. More interestingly, however, is that the formulas we derive will apply to a much wider class of possible AMMs. Whereas the idea above was described in the context of concentrated liquidity, we will instead find that our framework for internal fees applies in a broader setting with a minimal number of assumptions.

We now explicitly derive our general internal fee structure (GIFS). We must first begin by establishing an appropriately general setting, and to motivate this we make two initial observations. First, in a constant product pool, the coordinates (x, y) describing the amount of assets in the pool are given by the expressions

$$x(p) = L/\sqrt{p} \tag{1}$$
$$y(p) = L\sqrt{p} \tag{2}$$

where L is the total liquidity factor, and p is the current price. We think of L as a constant (more or less), as opposed to the ever fluctuating price, and thus we choose to write x and y explicitly as functions of p. Similarly, for a concentrated liquidity pool with a collection of active ticks $\{p_i\}$ and a net liquidity factor L_i for each interval, the total amount of assets in the pool is given by summing over all intervals:

$$x(p) = \sum_i L_i\Big(1/\sqrt{[p]_{p_i}^{p_{i+1}}} - 1/\sqrt{p_{i+1}}\Big) \tag{3}$$
$$y(p) = \sum_i L_i\Big(\sqrt{[p]_{p_i}^{p_{i+1}}} - \sqrt{p_i}\Big) \tag{4}$$

where the *box* function $[\cdot]$,

$$[t]_a^b := \left\{ \begin{array}{cc} a & \text{if } t < a \\ t & \text{if } a \le t \le b \\ b & \text{if } b < t \end{array} \right\}. \tag{5}$$

is useful to keep (3)-(4) compact and readable.

Inspired by these equations, we will now define our general setting. We suppose we have a protocol such that the total amount of tokens (x, y) are given as deterministic functions of the price p, with some parameter set $\{\alpha\}$:

$$x = x\big(p; \{\alpha\}\big) \tag{6}$$
$$y = y\big(p; \{\alpha\}\big) \tag{7}$$

We will be loose with how we write the parameters. Sometimes we may only write one of them, sometimes we omit them altogether (it should be clear from context).

Moreover, we assume that the total pool liquidity is a simple superposition over the collection of LPs, so that each LP has their own liquidity function, which we indicate with a subscript n for the n^{th} LP:

$$x_n = x_n\big(p; \{\alpha_n\}\big) \tag{8}$$
$$y_n = y_n\big(p; \{\alpha_n\}\big) \tag{9}$$

The superposition assumption means that the total amount of each asset is simply a sum over all LPs:

$$x = \sum_n x_n = \sum_n x_n\big(p; \{\alpha_n\}\big) = x\big(p; \{\alpha\}\big) \tag{10}$$
$$y = \sum_n y_n = \sum_n y_n\big(p; \{\alpha_n\}\big) = y\big(p; \{\alpha\}\big) \tag{11}$$

Of particular importance, we will assume that among the set of generic parameters $\{\alpha\}$, there exists a liquidity scaling parameter, which we denote by L, such that the functions are *linear* with respect to that parameter:

$$x_n\big(p; \eta L_n\big) = \eta\, x_n\big(p; L_n\big) \tag{12}$$
$$y_n\big(p; \eta L_n\big) = \eta\, y_n\big(p; L_n\big) \tag{13}$$

The liquidity factor L will be the only parameter of relevance for us in this discussion, but we nevertheless write (6)-(7) with a more abstract set of possible parameters $\{\alpha\}$ to allow for additional information (like, for example, the choice of price range $[p_a, p_b]$ that an LP chooses in concentrated liquidity).

Lastly, we will assume that the function $x(p)$ is monotonically *decreasing* with p, while the function $y(p)$ is monotonically *increasing* with p, and both are non-negative. Other than this, we make no further assumptions about the form of these functions. One can check that (1)-(2) and (3)-(4) both qualify.

When it comes to fees, we assume that our functions described above *do not* factor in fees. Instead, we impose a fee structure *on top* of these functions. We start with a fee factor $\varphi < 1$ that measures the fraction of the trade taken as a fee (relative to the conventional parameter γ, we work with φ = 1-γ). While fees are typically taken from the *incoming* quantity, we will split our parameter φ into a convex sum of both an *incoming* factor φ_i and an *outgoing* factor φ_o, using a parameter θ:

$$\varphi_i := \big(\theta\big)\varphi, \quad \varphi_o := \big(1\text{-}\theta\big)\varphi \quad (0 \le \theta \le 1) \tag{14}$$

With this parameter θ, we can (in theory) take a bit of the fee from the incoming quantity and a bit of the fee from the outgoing quantity, but still maintain an overall fee size of φ. For example, with a fee size of 0.3%, we might take 0.2% of the incoming quantity and 0.1% of the outgoing quantity.

Now, the functions $x(p)$ and $y(p)$ trace out a 'bonding' curve in the x/y plane as we vary over the prices p. Because these functions do not take fees into account, then any attempt to store the fees *internally* in the pool will necessarily deviate from this bonding curve (as in Figures 2 and 3). An elegant way to deal with this is to use the extra degree of freedom in our choice of θ to ensure that this deviation from the curve is *equivalent* to a simple scaling of the liquidity parameter L by some factor η (as in Figure 3). This allows us to implement our *fee-by-scaling* idea while still maintaining a well defined fee parameter φ. The math will be a bit complicated, but we can split it up into three distinct steps:

1. Find a formula for this required value of θ.
2. Find a computationally cheap approximation for it.
3. Redefine our approach, taking the approximation as *definition*, and find the resulting formula for η.

Before we begin our derivation, we will make few convenient choices of notation. First, we will omit the subscript n pertaining to the n^{th} LP, because our results could be applied at an individual LP level *or* in the aggregate, and for brevity we assume the latter. Second, for any given trade, our formulas will need to distinguish between the *incoming* quantity and the *outgoing* quantity. However, whether these quantities correspond to the x tokens or the y tokens depends on whether the price is moving down or up, respectively. In order to treat both cases at once, we adopt the following conventions to simplify our discussion; we denote the incoming and outgoing quantities by $(\Delta t_i, \Delta t_o)$, and the corresponding token values (before the trade) by (t_i, t_o). Explicitly, for a trade going from price p to p', we define

	if $\big(p' < p\big)$	if $\big(p' > p\big)$
t_i	$x(p)$	$y(p)$
t_o	$y(p)$	$x(p)$
Δt_i	$x(p') - x(p)$	$y(p') - y(p)$
Δt_o	$\lvert y(p') - y(p)\rvert$	$\lvert x(p') - x(p)\rvert$

We note that we take the absolute value for the outgoing quantity, because it is by definition negative, but for our formulas it will be cleaner to treat Δt_o as a positive quantity and simply attach a minus sign, when necessary.

Step 1: Computing θ

Now, according to our original assumptions, the quantities Δt_i and Δt_o are what would come in and out of the pool *in the absence of fees*. In this case, the amount of each asset in the pool would update according to

$$t_i \to t_i + \Delta t_i \tag{15}$$

$$t_o \to t_o - \Delta t_o \tag{16}$$

However, let us now consider how we can take fees and store them internally. If we take a fraction φ_i from the incoming quantity as a fee, then of course the remaining fraction 1-φ_i proceeds into the pool. This means that the trader must pay in the quantity $\Delta t_i/(1\text{-}\varphi_i)$, for then the amount going into the trade calculation would be $(1\text{-}\varphi_i)\big[\Delta t_i/(1\text{-}\varphi_i)\big] = \Delta t_i$, as it needs to be to move from p to p'. Moreover, when taking the fee from the outgoing quantity, a fraction φ_o must be removed from the asset Δt_o coming out of the pool, and thus the amount coming out of the pool is actually only $\Delta t_o(1\text{-}\varphi_o)$. Altogether then, the total change in the assets are given by

$$t_i \to t_i + \Delta t_i/(1\text{-}\varphi_i) \tag{17}$$

$$t_o \to t_o - \Delta t_o(1\text{-}\varphi_o) \tag{18}$$

Our goal, then, is to choose a value of θ (which appears implicitly in φ_i and φ_o above) such that the moves in (17)-(18) are *equivalent* to a two step sequence consisting of the *non-fee* moves of (15)-(16), followed by a scaling of the liquidity factor L up by some factor η (*fee-by-scaling*). But, from expressions (12)-(13), we see that scaling the liquidity is equivalent to scaling the coordinates (x, y). Thus, the preceding condition we wish to establish can be expressed as the following:

$$t_i + \Delta t_i/(1\text{-}\varphi_i) = \eta\Big(t_i + \Delta t_i\Big) \tag{19}$$

$$t_o - \Delta t_o(1\text{-}\varphi_o) = \eta\Big(t_o - \Delta t_o\Big) \tag{20}$$

This is illustrated in Figure 4 below. The image is a bit dense, but a careful examination may behoove the reader who desires a firm understanding. For the sake of illustration, we arbitrarily chose the case $p' < p$.

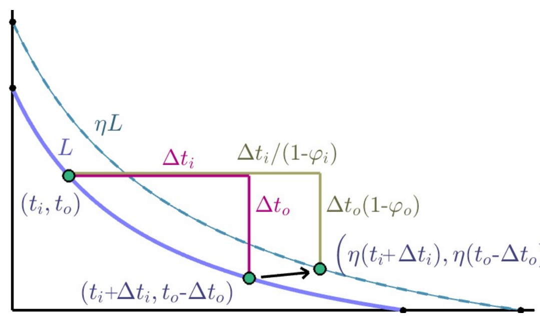


Figure 4. Incoming/outgoing fees, equivalent to scaling

Let us be careful to articulate the flow of information that gives rise to the equations (19)-(20). First, from a current price p and current asset quantities (t_i, t_o), a trader chooses a target price p'. With this price, the corresponding incoming/outgoing quantities (without fees) can then be computed as $(\Delta t_i, \Delta t_o)$. Hence, there are only two undetermined quantities in (19)-(20), being the scale factor of η, as well as the parameter θ appearing implicitly in φ_i and φ_o. Thus, we can think of (19)-(20) as a 2×2 system of equations in the unknowns (θ, η). If we divide (20) by (19), we eliminate η, and restoring the presence of θ via (14), this gives us

$$\frac{t_i + \Delta t_i/\big(1 - \theta\varphi\big)}{t_o - \Delta t_o\big(1 - (1\text{-}\theta)\varphi\big)} = \frac{t_i + \Delta t_i}{t_o - \Delta t_o} \tag{21}$$

This can be rearranged into a quadratic expression for θ:

$$0 = \Big[\varphi\Big]\theta^2 - \left[\frac{(t_i/\Delta t_i) + (t_o/\Delta t_o)}{(t_i/\Delta t_i) + 1} + \varphi\right]\theta + 1 \tag{22}$$

For the sake of brevity, we will define the following:

$$\delta := \frac{(t_i/\Delta t_i) + (t_o/\Delta t_o)}{(t_i/\Delta t_i) + 1}. \tag{23}$$

Then our quadratic expression becomes

$$0 = \big[\varphi\big]\theta^2 - \big[\delta + \varphi\big]\theta + 1. \tag{24}$$

If we define the quantity β by

$$\beta = \big(\delta + \varphi\big)/(2\varphi) \tag{25}$$

then the solution(s) to the quadratic equation (24) are given by

$$\theta = \beta \pm \sqrt{\beta^2 - 1/\varphi}. \tag{26}$$

It will be shown that we necessarily have $1/\varphi \le \beta^2$, so that the solutions to (26) are real. Moreover, it will also be shown that if we choose the *minus* solution of (26),

$$\theta = \beta - \sqrt{\beta^2 - 1/\varphi}, \tag{27}$$

then it will necessarily be the case that $0 < \theta \le 1$, which is fortunate since θ is meant to appear in a convex sum (14). Furthermore, it can also be shown that there exists a good approximation for (27) that avoids using square roots (which are computationally expensive) given by

$$\theta \approx \frac{\delta}{\delta^2 + (\delta - 1)\varphi}. \tag{28}$$

These claims will all be derived in due time, but to avoid losing our current momentum, let us continue on for the moment. Having found a value of θ, we can then find η using either (19) or (20). Choosing (19), we solve for η:

$$\eta = \frac{t_i + \Delta t_i/(1\text{-}\varphi_i)}{t_i + \Delta t_i} = 1 + \frac{\varphi_i/(1\text{-}\varphi_i)}{(t_i/\Delta t_i) + 1}. \tag{29}$$

Now we may carry out the trade by updating the asset amounts and the liquidity parameter:

$$t_i \to t_i + \Delta t_i/(1\text{-}\varphi_i) \tag{30}$$
$$t_o \to t_o - \Delta t_o(1\text{-}\varphi_o) \tag{31}$$
$$L \to \eta L \tag{32}$$

As we are going to develop this further still, this may be a good point to stop and make a cumulative summary of what we have achieved so far.

- For a trade from $p \to p'$, we define:

	if $(p' < p)$	if $(p' > p)$
t_i	$x(p)$	$y(p)$
t_o	$y(p)$	$x(p)$
Δt_i	$x(p') - x(p)$	$y(p') - y(p)$
Δt_o	$\lvert y(p') - y(p)\rvert$	$\lvert x(p') - x(p)\rvert$

- Then we compute the following sequence:

$$\delta = \frac{t_i/\Delta t_i + t_o/\Delta t_o}{t_i/\Delta t_i + 1}, \quad \beta = (\delta + \varphi)/(2\varphi)$$
$$\theta = \beta - \sqrt{\beta^2 - 1/\varphi}, \quad \varphi_i = \theta\varphi, \quad \varphi_o = (1\text{-}\theta)\varphi$$
$$\eta = 1 + \Big[\varphi_i/(1\text{-}\varphi_i)\Big]/\Big[t_i/\Delta t_i + 1\Big]$$

- Finally, we update:

$$t^i \to t^i + \Delta^i/(1\text{-}\varphi_i)$$
$$t^o \to t^o - \Delta^o(1\text{-}\varphi_o)$$
$$L \to \eta L$$

With this, we achieve the goal of taking both incoming and outgoing fees in a way equivalent to *fee-by-scaling*.

Before we proceed, we should confirm the claims made about the solution in (27). First, we note that from the non-negativity assumption on our liquidity functions, we cannot withdraw more of an asset than there exist in the pool, and so we can say that $\Delta t_o \le t_o$. In other words,

$$1 \le t_o/\Delta t_o. \tag{33}$$

Adding $(t_i/\Delta t_i)$ to both sides gives us

$$t_i/\Delta t_i + 1 \le t_i/\Delta t_i + t_o/\Delta t_o, \tag{34}$$

and upon dividing both sides by $t_i/\Delta t_i + 1$ we see our friend δ from definition (23) emerge:

$$1 \le \frac{t_i/\Delta t_i + t_o/\Delta t_o}{t_i/\Delta t_i + 1} = \delta. \tag{35}$$

Inequality (35) will be an important result for later:

$$1 \le \delta. \tag{36}$$

This of course means that $0 \le \delta - 1$ and so we may write

$$-(1 - \sqrt{\varphi})^2 \le \delta - 1. \tag{37}$$

But (37) simplifies to $2\sqrt{\varphi} \le \delta + \varphi$. Dividing both sides by 2φ, we find β from definition (25) emerge:

$$1/\sqrt{\varphi} \le (\delta + \varphi)/(2\varphi) = \beta \tag{38}$$

Squaring both sides of (38), we have

$$1/\varphi \le \beta^2. \tag{39}$$

Thus, for our θ solutions given in (26), we see that the solutions will be *real*, as promised.

Now we proceed to show that $0 < \theta \le 1$ when θ is given by (27), i.e. the *minus* solution. First, because $0 < 1/\varphi$, then we can say that $\beta^2 - 1/\varphi < \beta^2$, and taking the positive square root, we find $\sqrt{\beta^2 - 1/\varphi} < \beta$. But this then means that we have $0 < \beta - \sqrt{\beta^2 - 1/\varphi}$, which confirms that $0 < \theta$. Next, beginning with the observation that $1 \le \delta$ in (36), we can write $1 + \varphi \le \delta + \varphi$. Dividing both sides by φ gives

$$1/\varphi + 1 \le (\delta + \varphi)/\varphi = 2\beta \tag{40}$$

Squaring both sides of (40), we obtain $(1/\varphi + 1)^2 \le 4\beta^2$. Subtracting $4/\varphi$ from both sides then yields

$$(1/\varphi - 1)^2 \le 4(\beta^2 - 1/\varphi). \tag{41}$$

Now, because $\varphi < 1$, then $1/\varphi$ is greater than 1, so when we take the *positive square root* of (41), we will have

$$(1/\varphi - 1) \le 2\sqrt{\beta^2 - 1/\varphi}. \tag{42}$$

But by adding β^2 to both sides and rearranging, we find

$$\beta^2 \le (\beta^2 - 1/\varphi) + 2\sqrt{\beta^2 - 1/\varphi} + 1. \tag{43}$$

The right hand side of (43) is a perfect square, being $\left(\sqrt{\beta^2 - 1/\varphi} + 1\right)^2$, and so taking the positive square root of both sides of (43), we obtain $\beta \le \sqrt{\beta^2 - 1/\varphi} + 1$. Thus

$$\beta - \sqrt{\beta^2 - 1/\varphi} \le 1. \tag{44}$$

In other words, we have $\theta \le 1$. Combined with our previous observation, we conclude that $0 < \theta \le 1$.

Step 2: Approximating θ

As we mentioned in (28), we have computationally cheaper approximation for our formula (27) for θ. In this section, we derive this approximation. We will begin by modifying expression (27):

$$\begin{aligned}\theta &= \beta - \sqrt{\beta^2 - 1/\varphi} \\ &= \beta\Big(1 - \sqrt{1 - 1/(\varphi\beta^2)}\Big) \\ &= \frac{\delta+\varphi}{2\varphi}\left(1 - \sqrt{1 - \frac{4\varphi}{(\delta+\varphi)^2}}\right) \end{aligned} \tag{45}$$

With an eye towards using the binomial approximation $(1+\epsilon)^p \approx 1 + p\epsilon$ (for $\epsilon \ll 1$), we need to investigate the quantity $4\varphi/(\delta+\varphi)^2$ appearing in (45). For this we will make our first of two important assumptions (this being an approximation, after all). We begin with the following extremely reasonable statement:

$$\varphi \ll 1 \;\; \textbf{(Assumption 1)} \tag{46}$$

As a consequence, we use the inequality $1 \leq \delta$ established in (36) to write $\varphi \ll \delta$. Moreover, this gives us $0 \ll \delta - \varphi$. Being far above zero, we should be able to square the quantity $(\delta - \varphi)$ and still remain far above zero, obtaining $0 \ll (\delta-\varphi)^2$. Then, if we add the positive quantity $4\varphi(\delta - 1)$ to the right side, the statement sill holds, giving us $0 \ll (\delta-\varphi)^2 + 4\varphi(\delta-1)$. However, the right side simplifies to $0 \ll (\delta+\varphi)^2 - 4\varphi$, and so we can rearrange to obtain $4\varphi \ll (\delta+\varphi)^2$. Dividing through, we obtain

$$\frac{4\varphi}{(\delta+\varphi)^2} \ll 1. \tag{47}$$

Thus, we feel confident in applying the binomial approximation to (45):

$$\begin{aligned}\theta &= \frac{\delta+\varphi}{2\varphi}\left(1 - \sqrt{1 - \frac{4\varphi}{(\delta+\varphi)^2}}\right) \\ &\approx \frac{\delta+\varphi}{2\varphi}\left(1 - \left(1 - \frac{1}{2}\left(\frac{4\varphi}{(\delta+\varphi)^2}\right)\right)\right) \\ &= \frac{1}{\delta+\varphi}\end{aligned} \tag{48}$$

This gives us an initial approximation for θ, but we will actually take it a few steps further. However, to avoid confusion, let us temporarily distinguish this approximation from the real thing by denoting it with $\hat{\theta}$:

$$\hat{\theta} := 1/(\delta+\varphi). \tag{49}$$

In fact, let us momentarily think of our formulas for θ as *functions* of φ. In other words, for fixed δ, let $\theta(\varphi)$ be the function defined by (45) and $\hat{\theta}(\varphi)$ be defined by (49). Now, to improve on our approximation, let us ask the following question; when trying to evaluate $\theta(\varphi)$, is there a *better* fee factor φ' that we could plug into our approximation $\hat{\theta}$ such that the output is closer to what we want? In other words, we look for a φ' such that

$$\theta(\varphi) = \hat{\theta}(\varphi'). \tag{50}$$

To investigate this, let us play with some relationships. First, the *true* value of θ satisfies the quadratic equation in (24), namely $0 = \big(\varphi\big)\theta^2 - \big(\delta+\varphi\big)\theta + 1$ This is linear in φ, and so we can easily solve for it:

$$\varphi = \frac{1-\delta\theta}{\theta(1-\theta)} \tag{51}$$

But expression (51) is simply the relationship giving φ in terms of θ, meaning it is the *inverse* of the function $\theta(\varphi)$, i.e. $\theta^{-1}(\cdot)$. Applying it to both sides of (50) gives us

$$\varphi = \theta^{-1}\Big(\hat{\theta}(\varphi')\Big) \tag{52}$$

We can make this explicit by directly substituting our expression for $\hat{\theta}$ from (49) into θ^{-1} from (51):

$$\begin{aligned}\varphi &= \frac{1-\delta\,\hat{\theta}(\varphi')}{\hat{\theta}(\varphi')\big(1-\hat{\theta}(\varphi')\big)} \\ &= \frac{1-\delta\left(\frac{1}{\delta+\varphi'}\right)}{\left(\frac{1}{\delta+\varphi'}\right)\left(1-\left(\frac{1}{\delta+\varphi'}\right)\right)} \\ &= \frac{\varphi'(\delta+\varphi')}{(\delta+\varphi'-1)}\end{aligned} \tag{53}$$

To set ourselves up for another binomial approximation, let's rearrange (53) a little bit more:

$$\begin{aligned}\varphi &= \frac{\varphi'(\delta+\varphi')}{(\delta+\varphi'-1)} \\ &= \varphi' + \varphi'\left(\frac{1}{\delta-1+\varphi'}\right) \\ &= \varphi' + \frac{\varphi'}{\delta-1}\left(\frac{1}{1+\frac{\varphi'}{\delta-1}}\right).\end{aligned} \tag{54}$$

It is time to invoke another assumption. We will assume that the exchange rate $\Delta t_i/\Delta t_o$ is similar to the asset ratio t^i/t^o (as it would be for the constant product AMM, for example):

$$t_i/t_o \approx \Delta t_i/\Delta t_o \;\; \textbf{(Assumption 2)} \tag{55}$$

We note that this assumption should hold *before or after* the trade. After all, the exchange rate between two prices is independent of the direction of the trade. Specifically, this means that according to assumption 2, we should *also* have

$$(t_i+\Delta t_i)/(t_o-\Delta t_o) \approx \Delta t_i/\Delta t_o. \tag{56}$$

Cross multiplying, this gives us

$$\frac{t^i/\Delta^i + 1}{t^o/\Delta^o - 1} \approx 1. \tag{57}$$

By good fortune, we recognize the left side of (57) to be the reciprocal of $(\delta - 1)$:

$$\begin{aligned}\delta - 1 &= \frac{t_i/\Delta t_i + t_o/\Delta t_o}{t_i/\Delta t_i + 1} - 1 \\ &= \frac{\big(t_i/\Delta t_i + t_o/\Delta t_o\big) - \big(t_i/\Delta t_i + 1\big)}{t_i/\Delta t_i + 1} \\ &= \frac{t_o/\Delta t_o - 1}{t_i/\Delta t_i + 1}. \end{aligned} \tag{58}$$

Thus, (57) tells us that

$$\frac{1}{\delta - 1} \approx 1 \tag{59}$$

If we multiply (59) by $\varphi' \ll 1$, then we should have $\varphi'/(\delta - 1) \ll 1$ as well, and so we confidently apply the binomial approximation to (54):

$$\begin{aligned}\varphi &= \varphi' + \frac{\varphi'}{\delta - 1}\left(\frac{1}{1 + \frac{\varphi'}{\delta-1}}\right) \\ &\approx \varphi' + \frac{\varphi'}{\delta - 1}\Big(1 - \frac{\varphi'}{\delta - 1}\Big) \\ &= \left(\frac{\delta}{\delta - 1}\right)\varphi' + O\Big[(\varphi')^2\Big]. \end{aligned} \tag{60}$$

Thus, to first order in φ', we have $\varphi \approx \left(\frac{\delta}{\delta-1}\right)\varphi'$, or in other words,

$$\varphi' \approx \left(\frac{\delta - 1}{\delta}\right)\varphi \tag{61}$$

Now, recalling that our original goal is to approximate the value $\theta(\varphi)$, we then plug (61) into (50):

$$\begin{aligned}\theta(\varphi) &= \hat{\theta}(\varphi') \\ &= \left(\frac{1}{\delta + \varphi'}\right) \\ &\approx \left(\frac{1}{\delta + \left(\frac{\delta-1}{\delta}\right)\varphi}\right) \\ &= \frac{\delta}{\Big(\delta^2 + (\delta - 1)\varphi\Big)} \end{aligned} \tag{62}$$

Thus, we confirm the expression in (28) for our glorious approximation for evaluating θ:

$$\theta \approx \frac{\delta}{\delta^2 + (\delta - 1)\varphi}. \tag{63}$$

It is glorious, because it does not require any square roots. Moreover, numerical testing shows that (63) agrees *very* closely with (27) for a wide variety of φ and δ values. Nevertheless, there is an obvious issue we must now address; in what sense are we even *allowed* to make approximations? In the realm of AMMs, calculations must be *exact* (up to a tolerance) or else the transaction can not be executed. So of what use, then, is (63)?

Step 3: Redefining θ and η

To address this question of the allowability of our approximation, we must now adopt a change of perspective. In order to explain this change, let us first consider order of operations, as it currently stands:

1. We have a static fee parameter φ.
2. For a given trade, our formula produces a particular value θ. This value splits our fee parameter φ into two components φ_i and φ_o, with which we take a fee from both the incoming and outgoing quantities.
3. It just so happens to be the case that this particular value of θ guarantees that what are doing is equivalent to *fee-by-scaling*, by a factor η.

Now, if we insist on using our approximation to evaluate θ in step (2) above, then the *equivalency* claimed in step (3) is no longer legitimate. However, consider the following reframing of this sequence of steps:

1. We have a static fee parameter φ.
2. For a given trade, we use our approximation for θ and obtain the corresponding η value. We explicitly (and happily) do *fee-by-scaling* with this factor η.
3. There exists some *effective* fee factor $\widetilde{\varphi}$ that is *implicit* in our choice of η. Because of the use of our approximation formula, this effective fee factor will be close to our static *target* fee factor φ.

In other words, our change in perspective is that we are now demoting our fee parameter φ from its original position as the *actual* fee size. Instead, we can think of φ as a *target* fee factor, and its role in the approximation for θ is now just that of an input ingredient that allows us to compute a scale factor η, which we then use explicitly to do *fee-by-scaling*. Doing this implicitly establishes an effective fee parameter $\widetilde{\varphi}$, and as we will show, this will only end up differing from the target fee factor φ by an amount that is $< 1\%$.

To make this concrete, let us again suppose we are executing a trade from $p \to p'$, with $(t^i, t^o, \Delta t^i, \Delta t^o)$ defined as before. Then we have the following sequence:

- Using a target fee factor φ, we use approximation (63) for θ (which then gives us φ_i, φ_o from (14)).
- We then produce a scale factor η using (29).
- Using η, we explicitly do *fee-by-scaling*:

$$t_i \to \eta\Big(t_i + \Delta t_i\Big) \tag{64}$$

$$t_o \to \eta\Big(t_o - \Delta t_o\Big) \tag{65}$$

Our remaining goal, then, is twofold. We wish to understand what kind of effective fee factor $\widetilde{\varphi}$ is *implicit* in the steps (64)-(65). Moreover, in doing so, we will also find efficient formulas for actually executing (64)-(65).

Now, for *fee-by-scaling*, we have been talking about it as a trade *without fees*, followed by liquidity scaling. At the end of the transaction, however, there will simply be some change in the asset quantities. Let us denote these actual changes by $\underline{\Delta} t^i$ and $\underline{\Delta} t^o$, so that we may write

$$\eta\Big(t_i + \Delta t_i\Big) = t_i + \underline{\Delta} t_i \tag{66}$$

$$\eta\Big(t_o - \Delta t_o\Big) = t_o - \underline{\Delta} t_o \tag{67}$$

We rearrange (66)-(67) for these deltas:

$$\underline{\Delta} t_i = \eta \Delta t_i + (\eta\text{-}1) t_i \tag{68}$$

$$\underline{\Delta} t_o = \eta \Delta t_o - (\eta\text{-}1) t_o \tag{69}$$

We claimed that implicit in this *fee-by-scaling* is an effective fee factor $\widetilde{\varphi}$. Let us now investigate. Suppose such a $\widetilde{\varphi}$ exists, and suppose that it results in a value $\widetilde{\theta}$ after using our formula (27). In terms of these quantities, the actual changes in the assets will be given by our formulas from (19)-(20), which can now be written as:

$$\underline{\Delta} t_i = \Delta t_i / \big(1 - \widetilde{\theta}\widetilde{\varphi}\big) \tag{70}$$

$$\underline{\Delta} t_o = \Delta t_o \big(1 - (1\text{-}\widetilde{\theta})\widetilde{\varphi}\big) \tag{71}$$

Because of the fact that $(\Delta t_i, \Delta t_o)$ are determined beforehand, and $(\underline{\Delta} t_i, \underline{\Delta} t_o)$ are determined once we choose η, then the two equations (70)-(71) have only two unknowns $(\widetilde{\theta}, \widetilde{\varphi})$, and so their values should be determined as well. Our goal will be to find an insightful expression for $\widetilde{\varphi}$, so that we may understand what kind of fees we will effectively be imposing. Thus, let us begin by solving for $\widetilde{\varphi}$. We can rearrange (70)-(71) as follows:

$$\Delta t_i / \underline{\Delta} t_i = \big(1 - \widetilde{\theta}\widetilde{\varphi}\big) \tag{72}$$

$$\underline{\Delta} t_o / \Delta t_o = \big(1 + \widetilde{\theta}\widetilde{\varphi} - \widetilde{\varphi}\big) \tag{73}$$

Adding the two equations together eliminates the $\widetilde{\theta}$ term, and so we may solve for $\widetilde{\varphi}$:

$$\widetilde{\varphi} = \Big(1 - \Delta t_i / \underline{\Delta} t_i\Big) + \Big(1 - \underline{\Delta} t_o / \Delta t_o\Big) \tag{74}$$

Now, we would like to express $\widetilde{\varphi}$ in terms of the basic state quantities $(t_i, t_o, \Delta t_i, \Delta t_o)$. This is doable, but it will take us a minute. First, let us look at the quantities in parathenses. From (68)-(69), we can write

$$\underline{\Delta} t_i / \Delta t_i = \eta + (\eta\text{-}1) t_i / \Delta t_i \tag{75}$$

$$\underline{\Delta} t_o / \Delta t_o = \eta - (\eta\text{-}1) t_o / \Delta t_o \tag{76}$$

We can put these into (74) and simplify:

$$\begin{aligned}
\widetilde{\varphi} &= \Big(1 - \Delta t_i / \underline{\Delta} t_i\Big) + \Big(1 - \underline{\Delta} t_o / \Delta t_o\Big) \\
&= \left[1 - \frac{1}{\eta + (\eta\text{-}1) t_i / \Delta t_i}\right] + \Big[1 - \Big(\eta - (\eta\text{-}1) t_o / \Delta t_o\Big)\Big] \\
&= \frac{(\eta\text{-}1)(t_i / \Delta t_i + 1)}{1 + (\eta\text{-}1)(t_i / \Delta t_i + 1)} + (\eta\text{-}1)\big(t_o / \Delta t_o - 1\big)
\end{aligned} \tag{77}$$

To clean up (77), we will define two useful quantities:

$$\tau_i := t_i / \Delta t_i + 1 \tag{78}$$

$$\tau_o := t_o / \Delta t_o - 1 \tag{79}$$

Then (77) can be written as

$$\begin{aligned}
\widetilde{\varphi} &= \frac{(\eta\text{-}1)\tau_i}{1 + (\eta\text{-}1)\tau_i} + (\eta\text{-}1)\tau_o \\
&= (\eta\text{-}1)\left(\frac{\tau_i + \tau_o + (\eta\text{-}1)\tau_i\tau_o}{1 + (\eta\text{-}1)\tau_i}\right)
\end{aligned} \tag{80}$$

Because it will be convenient, we pull out a factor of τ_i:

$$\widetilde{\varphi} = (\eta\text{-}1)\tau_i \left(\frac{1 + \tau_o/\tau_i + (\eta\text{-}1)\tau_i(\tau_o/\tau_i)}{1 + (\eta\text{-}1)\tau_i}\right) \tag{81}$$

Now, we observe that with our definitions of τ_i and τ_o, then δ defined in (23) satisfies

$$\delta = (\tau_i + \tau_o)/\tau_i = 1 + \tau_o/\tau_i, \tag{82}$$

and thus

$$(\delta - 1) = \tau_o/\tau_i. \tag{83}$$

Substituting this into (81), we have

$$\widetilde{\varphi} = (\eta\text{-}1)\tau_i \left(\frac{\delta + (\eta\text{-}1)\tau_i(\delta - 1)}{1 + (\eta\text{-}1)\tau_i}\right) \tag{84}$$

Because of this last manipulation, we notice that the expression $(\eta\text{-}1)\tau_i$ appears in several places. Using (29), and our definition of τ_i, this can be written as

$$(\eta\text{-}1)\tau_i = \varphi_i/(1\text{-}\varphi_i). \tag{85}$$

We therefore focus in on the term φ_i. First, using our current *definition* of θ (i.e. formula (63)), we multiply by φ according to (14) and obtain

$$\varphi_i = \frac{\delta\varphi}{\delta^2 + (\delta - 1)\varphi}. \tag{86}$$

Moreover, we can then express 1-φ^i as

$$\begin{aligned}
1\text{-}\varphi_i &= 1 - \frac{\delta\varphi}{\delta^2 + (\delta - 1)\varphi} \\
&= \frac{\delta^2 - \varphi}{\delta^2 + (\delta - 1)\varphi}.
\end{aligned} \tag{87}$$

Dividing (86) by (87), we get the following expression:

$$\varphi_i/(1\text{-}\varphi_i) = \frac{\delta\varphi}{\delta^2 - \varphi}. \tag{88}$$

Thus, using equations (85) and (88) transitively, we can now express $(\eta\text{-}1)\tau_i$ as $\delta\varphi/(\delta^2 - \varphi)$, and so (84) becomes

$$\widetilde{\varphi} = \left(\frac{\delta\varphi}{\delta^2 - \varphi}\right)\left(\frac{\delta + \left(\frac{\delta\varphi}{\delta^2-\varphi}\right)(\delta - 1)}{1 + \left(\frac{\delta\varphi}{\delta^2-\varphi}\right)}\right) \tag{89}$$

We can clear the denominators of (89) and simplify:

$$\widetilde{\varphi} = \frac{\delta^2\varphi\Big(\delta^2+\varphi(\delta-1)-\varphi\Big)}{(\delta^2-\varphi)\Big(\delta^2+\varphi(\delta-1)\Big)} = \frac{\delta^3\varphi(\delta-\varphi)}{(\delta^2-\varphi)\big(\delta^2-(\delta-1)\varphi\big)} \tag{90}$$

Let us lastly divide (90) by φ:

$$\frac{\widetilde{\varphi}}{\varphi} = \frac{\delta^3(\delta-\varphi)}{(\delta^2-\varphi)\big(\delta^2-(\delta-1)\varphi\big)} \tag{91}$$

The ratio between our *ideal* fee factor φ and our *actual* fee factor $\widetilde{\varphi}$ is now made explicit with expression (91). Immediately, we can evaluate a few limits:

$$\lim_{\delta\to\infty} \widetilde{\varphi}/\varphi = 1 \tag{92}$$

$$\lim_{\varphi\to 0} \widetilde{\varphi}/\varphi = 1 \tag{93}$$

Note that we don't need to consider the limit of $\delta \to 0$ because we already know that $1 \leq \delta$. Moreover, we don't need to consider the limit of $\varphi \to \infty$, or even $\varphi \to 1$, because the fee factor will always be taken to be small. Thus, (92)-(93) are the only relevant limits, and they are certainly encouraging, as we would like $\widetilde{\varphi}$ to be as close to φ as possible. We can get a more quantitative assessment by using a binomial expansion two more times. First, let's prep expression (91):

$$\frac{\widetilde{\varphi}}{\varphi} = \frac{\delta^3(\delta-\varphi)}{(\delta^2-\varphi)\big(\delta^2-(\delta-1)\varphi\big)} = \delta^3(\delta-\varphi)\left[\frac{1}{\delta^2\left(1-\frac{\varphi}{\delta^2}\right)}\right]\left[\frac{1}{\delta^2\left(1-\frac{(\delta-1)\varphi}{\delta^2}\right)}\right] \tag{94}$$

Now, because we know that $1 \leq \delta$ and $\varphi \ll 1$, then the first factor in square brackets definitely deserves a binomial approximation. Moreover, we can also say that $\delta - 1 < \delta \leq \delta^2$, and so $(\delta-1)/\delta^2 < 1$. Again, combined with $\varphi \ll 1$, this justifies the binomial approximation on the second term in square brackets. Thus, after first cleaning up a bit, we have

$$\frac{\widetilde{\varphi}}{\varphi} = \left(1-\frac{\varphi}{\delta}\right)\left[\frac{1}{\left(1-\frac{\varphi}{\delta^2}\right)}\right]\left[\frac{1}{\left(1-\frac{(\delta-1)\varphi}{\delta^2}\right)}\right] \approx \left(1-\frac{\varphi}{\delta}\right)\left(1+\frac{\varphi}{\delta^2}\right)\left(1+\frac{(\delta-1)}{\delta^2}\varphi\right) \tag{95}$$

Collecting powers of φ, we have

$$\frac{\widetilde{\varphi}}{\varphi} = 1 - \left(\frac{1+\delta(\delta-1)}{\delta^4}\right)\left[\varphi^2\right] + O\left[\varphi^3\right] \tag{96}$$

This is also encouraging; the amount by which $\widetilde{\varphi}/\varphi$ differs from one is, at worst, *quadatic* in φ, which itself is small.

As a rough estimate, recall **assumption 2**, which states $t^i/\Delta t^i \approx t^o/\Delta t^o$, as well as an often reasonable assumption that $\Delta t^i \ll t^i$, with which we can get a reasonable numeric estimate out of our definition of δ:

$$\delta = \frac{t_i/\Delta t_i + t_o/\Delta t_o}{t_i/\Delta t_i + 1} \approx \frac{2(t_i/\Delta t_i)}{t_i/\Delta t_i + 1} = \frac{2}{1+\Delta t_i/t_i} \approx 2 \tag{97}$$

Then (96) becomes

$$\frac{\widetilde{\varphi}}{\varphi} \approx 1 - \frac{3}{16}\left[\varphi^2\right] \tag{98}$$

Using a (relatively large) fee size of $\varphi = 0.01$ (or 1%), then this says

$$\frac{\widetilde{\varphi}}{\varphi} \approx 1 - \left(1.875 \times 10^{-5}\right) = 0.99998125 \tag{99}$$

Now that we feel justifiably confident that our effective fee factor $\widetilde{\varphi}$ will be sufficient, we are almost ready to conclude. Before we do, however, we can take advantage of all the tedious preceding calculations in order to find a computationally cheap formula for computing η. Using a combination of both (85), (88) and (82), we proceed:

$$(\eta\text{-}1) = \frac{1}{\tau_i}\frac{\delta\varphi}{\delta^2-\varphi} = \frac{1}{\tau_i}\frac{(1+\tau_o/\tau_i)\varphi}{(1+\tau_o/\tau_i)^2-\varphi} = \frac{(\tau_i+\tau_o)\varphi}{(\tau_i+\tau_o)^2-(\tau_i)^2\varphi} = \frac{(t_i/\Delta t_i + t_o/\Delta t_o)\varphi}{(t_i/\Delta t_i + t_o/\Delta t_o)^2 - (t_i/\Delta t_i + 1)^2\varphi}. \tag{100}$$

With an eye towards minimizing the number of divisions we must compute, we multiply and divide expression (100) by a factor of $(\Delta t_i \Delta t_o)^2$, obtaining

$$(\eta\text{-}1) = \frac{\big[\Delta t_i\Delta t_o\big]\big[t_i\Delta t_o + t_o\Delta t_i\big]\varphi}{\big[t_i\Delta t_o + t_o\Delta t_i\big]^2 - \big[t_i\Delta t_o + \Delta t_i\Delta t_o\big]^2\varphi}. \tag{101}$$

Expression (101) may be the most arithmetically simple expression for η that we can hope for. Specifically, given the quantities $(t^i, t^o, \Delta t^i, \Delta t^o)$, we would first compute

$$a := t_i\Delta t_o, \quad b := t_o\Delta t_i, \quad c := \Delta t_i\Delta t_o \tag{102}$$

and then η can be expressed as

$$\eta = 1 + \frac{c(a+b)\varphi}{(a+b)^2-(a+c)^2\varphi} \tag{103}$$

This requires 4 additions, 8 multiplications, 1 division, and *no square roots.* More importantly, we see that with this recipe we can now completely bypass the unnecessary calculation of θ or φ_i. Instead, we just go straight to the calculation of η, followed by the calculation of the deltas $\underline{\Delta}t^i$ and $\underline{\Delta}t_o$ according to (68)-(69).

Section 3: Algorithm Summary & Discussion

Having derived all of our results, let us now give a complete and concise summary for our general internal fee structure (GIFS). Our setting is a liquidity pool containing two asset types x and y, and we assume the existence of two functions $x(p)$ and $y(p)$ giving the total amount of each asset present in the pool, as functions of the current price p. We make few assumptions on the nature of these functions, other than

- We assume that $x(p)$ is decreasing with respect to the price p, while $y(p)$ increases with respect to p, and both are always non-negative.
- There exists a liquidity scale parameter L, on which both x and y depend *linearly*.
- There is a fixed target fee factor φ.

(One notes that both *constant product* and *concentrated liquidity* fall within this set of assumptions, as demonstrated at the start of Section 2.) To impose a fee structure *on top* of these functions $x(p)$ and $y(p)$, we consider a trade that moves the price from $p \to p'$, and we have the following sequence of steps:

- We define the *incoming* and *outgoing* quantities (the current values t_i, t_o, and their deltas Δt^i, Δt^o):

	if $(p' < p)$	if $(p' > p)$
t^i	$x(p)$	$y(p)$
t^o	$y(p)$	$x(p)$
Δt^i	$x(p') - x(p)$	$y(p') - y(p)$
Δt^o	$y(p) - y(p')$	$x(p) - x(p')$

- We compute a scale factor η via the following:

$$\boxed{a := t^i \Delta t^o} \quad \boxed{b := t^o \Delta t^i} \quad \boxed{c := \Delta t^i \Delta t^o}$$

$$\boxed{\eta = 1 + \frac{c(a+b)\,\varphi}{(a+b)^2 - (a+c)^2\varphi}}$$

- We compute the actual deltas $(\underline{\Delta} t^i, \underline{\Delta} t^o)$ by:

$$\boxed{\underline{\Delta} t^i = \eta \Delta t^i + (\eta\text{-}1)t^i} \quad \boxed{\underline{\Delta} t^o = \eta \Delta t^o - (\eta\text{-}1)t^o}$$

- We can then update the asset amounts:

$$\boxed{t^i \to t^i + \underline{\Delta} t^i} \quad \boxed{t^o \to t^o - \underline{\Delta} t^o}$$

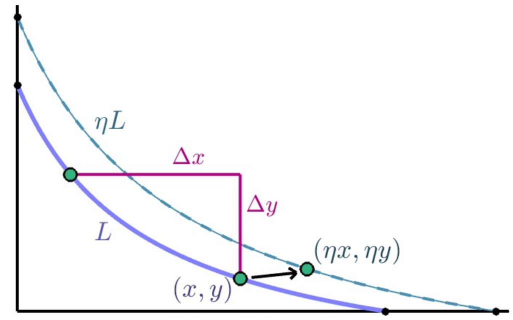
- This requires an update to the parameter L:

$$\boxed{L \to \eta L}$$

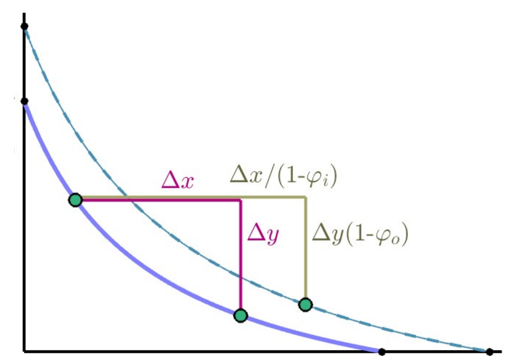
We call this algorithm the **GIFS** method.

We can interpret this procedure in one of two ways:

- We are executing the trade strictly according to the functions $x(p)$ and $y(p)$, as though there were *no fees*, then followed up with a scaling of the liquidity by an overall factor of η (what we call *fee-by-scaling*).



- Alternatively, we can think of this as though we are taking our fee factor φ and splitting it into two components φ_i and φ_o to be applied to the *incoming* and *outgoing* quantities. These satisfy $\varphi = \varphi_i + \varphi_o$, so for example, a fee size of 0.3% might be split into an incoming fee of 0.2% and an outgoing fee of 0.1%.



The only caveat here is that it is not *quite* our ideal fee factor φ that is being split, but rather an effective fee factor $\widetilde{\varphi}$. This value will vary from trade to trade, but generally we have the result

$$\boxed{\widetilde{\varphi}/\varphi = 1 - O\big[\varphi^2\big]}$$

so that it diverges from our ideal fee factor by an amount that is < 1% off from the target.

Through our hard earned formulas, we confirm that these two interpretations are **mathematically equivalent**.

The final question we address now is applications. In Section 1, we motivated our development in the context of a concentrated liquidity AMM. Indeed, this is the most obvious setting in which one would want to apply the GIFS. We leave the details of how this implementation would be done for another paper. Equally exciting, however, are some other AMM designs that we explore here and here, for which the GIFS method applies just as well.

Lastly, we note that it is also acceptable to let GIFS stand for the *Glorious* Internal Fee Strucure.