PDAs Accept Context-Free Languages

Theorem:

$$\left\{\begin{array}{c}\text{Context-Free}\\ \text{Languages}\\ \text{(Grammars)}\end{array}\right\} = \left\{\begin{array}{c}\text{Languages}\\ \text{Accepted by}\\ \text{PDAs}\end{array}\right\}$$

Proof - Step 1:

$$\left\{\begin{array}{c}\text{Context-Free}\\ \text{Languages}\\ \text{(Grammars)}\end{array}\right\} \subseteq \left\{\begin{array}{c}\text{Languages}\\ \text{Accepted by}\\ \text{PDAs}\end{array}\right\}$$

Convert any context-free grammar G to a PDA M with: $L(G) = L(M)$

Proof - Step 2:

$$\left\{\begin{array}{c}\text{Context-Free}\\ \text{Languages}\\ \text{(Grammars)}\end{array}\right\} \supseteq \left\{\begin{array}{c}\text{Languages}\\ \text{Accepted by}\\ \text{PDAs}\end{array}\right\}$$

Convert any PDA M to a context-free grammar G with: $L(G) = L(M)$

Proof - step 1

Convert

Context-Free Grammars
to
PDAs

Take an arbitrary context-free grammar G

We will convert G to a PDA M such that:

$$L(G) = L(M)$$

Conversion Procedure:

For each production in G

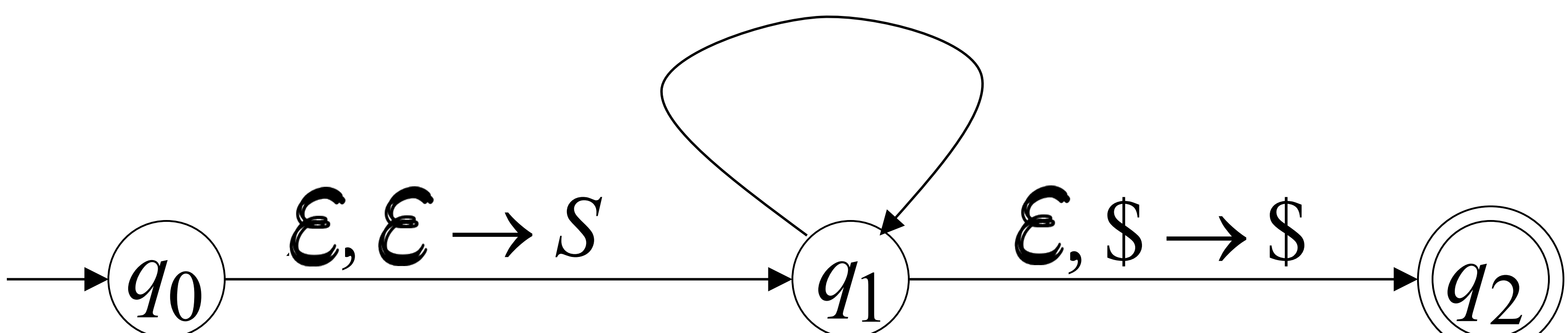
$$A \to w$$

For each terminal in G

$$a$$

Add transitions

$$\varepsilon, A \to w \qquad a, a \to \varepsilon$$



Example

Grammar

$S \to aSTb$

$S \to b$

$T \to Ta$

$T \to \varepsilon$

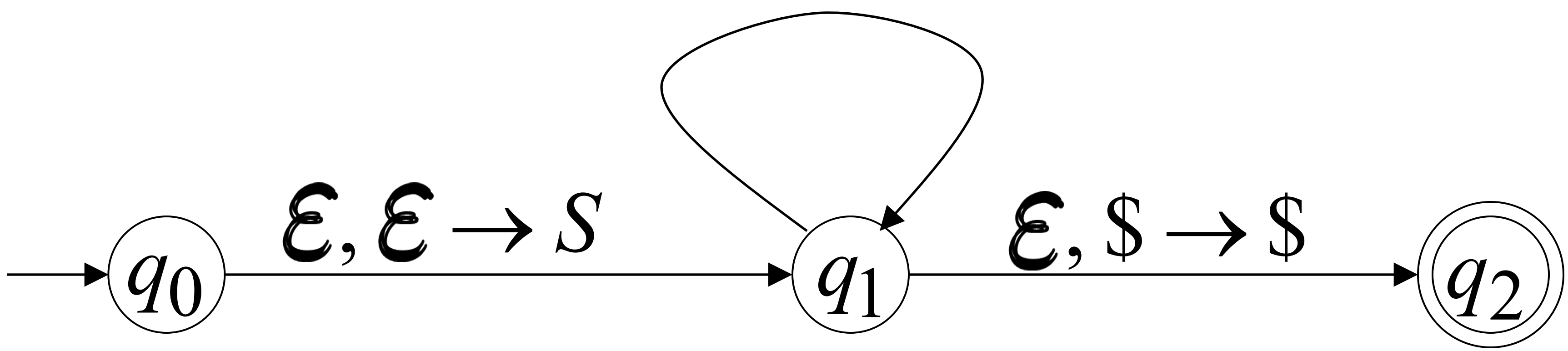
PDA

$\varepsilon, S \to aSTb$

$\varepsilon, S \to b$

$\varepsilon, T \to Ta$ $\qquad a, a \to \varepsilon$

$\varepsilon, T \to \varepsilon$ $\qquad b, b \to \varepsilon$



PDA simulates leftmost derivations

Grammar
Leftmost Derivation

S

$\Rightarrow \cdots$

$\Rightarrow \boxed{\sigma_1 \cdots \sigma_k} \boxed{X_1 \cdots X_m}$

$\Rightarrow \cdots$

$\Rightarrow \sigma_1 \cdots \sigma_k \sigma_{k+1} \cdots \sigma_n$

PDA Computation

$(q_0, \sigma_1 \cdots \sigma_k \sigma_{k+1} \cdots \sigma_n, \$)$

$\succ (q_1, \boxed{\sigma_1 \cdots \sigma_k} \sigma_{k+1} \cdots \sigma_n, S\$)$

$\succ \cdots$

$\succ (q_1, \sigma_{k+1} \cdots \sigma_n, \boxed{X_1 \cdots X_m} \$)$

$\succ \cdots$

$\succ (q_2, \varepsilon, \$)$

Scanned symbols

Stack contents

Grammar

Leftmost Derivation

Terminals

Leftmost variable

Variables or terminals

$$\Rightarrow \cdots$$

$$\Rightarrow xAy$$

$$\Rightarrow x\sigma_i \cdots \sigma_j Bzy$$

Production applied

$$A \rightarrow \underbrace{\sigma_i \cdots \sigma_j}_{\text{Terminals}} Bz$$

Variables or terminals

Variable

Grammar
Leftmost Derivation

PDA Computation

$\Rightarrow \cdots$

$\Rightarrow xAy$

$\Rightarrow x\sigma_i \cdots \sigma_j Bzy$

$\succ \cdots$

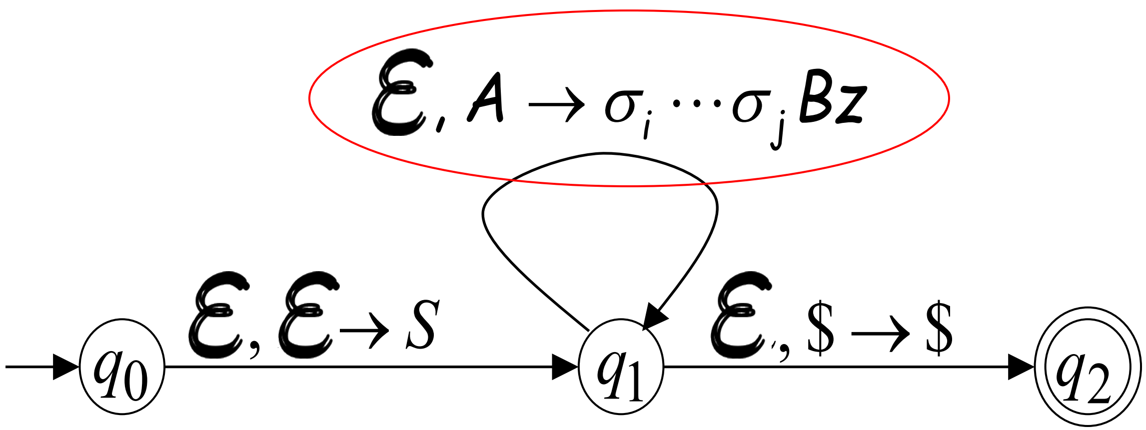
$\succ (q_1, \sigma_i \cdots \sigma_n, Ay\$)$

$\succ (q_1, \sigma_i \cdots \sigma_n, \sigma_i \cdots \sigma_j Bzy\$)$

Production applied

$A \rightarrow \sigma_i \cdots \sigma_j Bz$

Transition applied



Grammar

Leftmost Derivation

PDA Computation

$$\Rightarrow \cdots$$

$$\Rightarrow xAy$$

$$\Rightarrow x\sigma_i \cdots \sigma_j Bzy$$

$$\succ \cdots$$

$$\succ (q_1, \sigma_i \cdots \sigma_n, Ay\$)$$

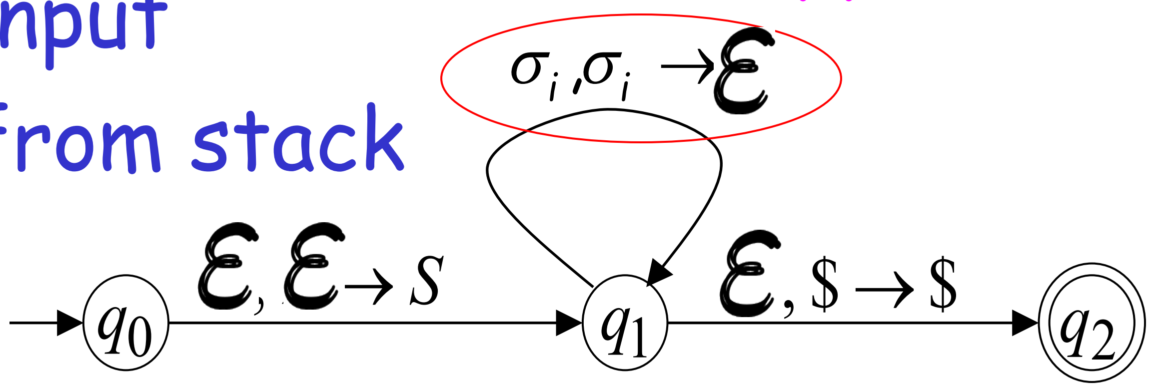
$$\succ (q_1, \sigma_i \cdots \sigma_n, \sigma_i \cdots \sigma_j Bzy\$)$$

$$\succ (q_1, \sigma_{i+1} \cdots \sigma_n, \sigma_{i+1} \cdots \sigma_j Bzy\$)$$

Transition applied

Read σ_i from input

and remove it from stack



Grammar Leftmost Derivation

$\Rightarrow \cdots$

$\Rightarrow xAy$

$\Rightarrow x\sigma_i \cdots \sigma_j Bzy$

PDA Computation

$\succ \cdots$

$\succ (q_1, \sigma_i \cdots \sigma_n, Ay\$)$

$\succ (q_1, \sigma_i \cdots \sigma_n, \sigma_i \cdots \sigma_j Bzy\$)$

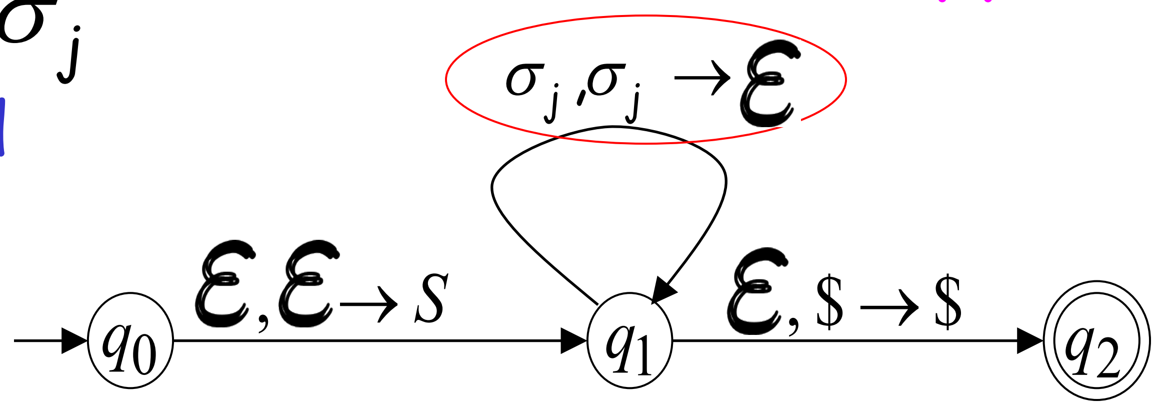
$\succ (q_1, \sigma_{i+1} \cdots \sigma_n, \sigma_{i+1} \cdots \sigma_j Bzy\$)$

$\succ \cdots$

$\succ (q_1, \sigma_{j+1} \cdots \sigma_n, Bzy\$)$

Last Transition applied

All symbols $\sigma_i \cdots \sigma_j$ have been removed from top of stack



The process repeats with the next leftmost variable

$$\Rightarrow \cdots$$

$$\Rightarrow xAy$$

$$\Rightarrow x\sigma_i \cdots \sigma_j Bzy$$

$$\Rightarrow x\sigma_i \cdots \sigma_j \sigma_{j+1} \cdots \sigma_k Cpzy$$

$$\succ \cdots$$

$$\succ (q_1, \sigma_{j+1} \cdots \sigma_n, Bzy\$)$$

$$\succ (q_1, \sigma_{j+1} \cdots \sigma_n, \sigma_{j+1} \cdots \sigma_k Cpzy\$)$$

$$\succ \cdots$$

$$\succ (q_1, \sigma_{k+1} \cdots \sigma_n, Cpzy\$)$$

Production applied

$$B \rightarrow \sigma_{j+1} \cdots \sigma_k Cp$$

And so on.....

Example:

Input

a	b	a	b

Time 0

$$\varepsilon, S \to aSTb$$

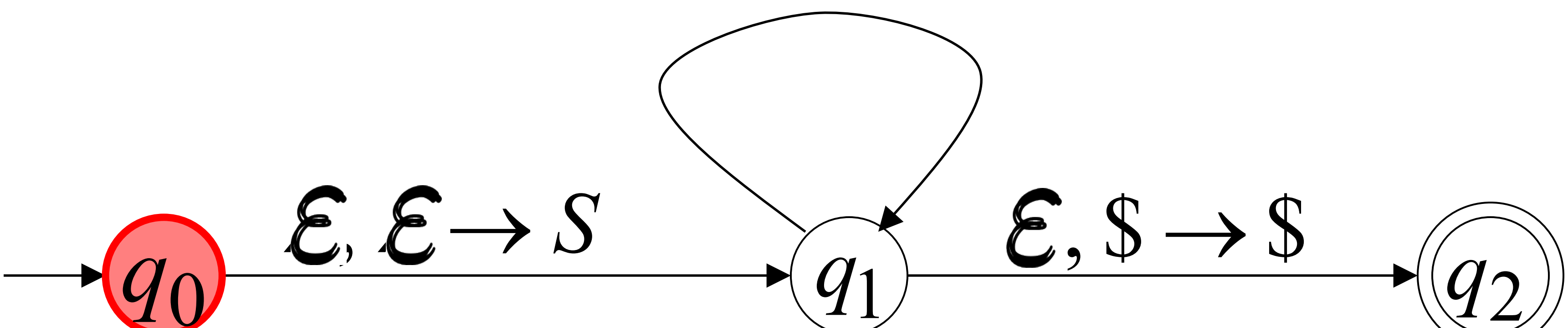
$$\varepsilon, S \to b$$

$$\varepsilon, T \to Ta \qquad a, a \to \varepsilon$$

$$\varepsilon, T \to \varepsilon \qquad b, b \to \varepsilon$$

$

Stack



Derivation: S

Input

a	b	a	b

Time 1

$\varepsilon, S \to aSTb$

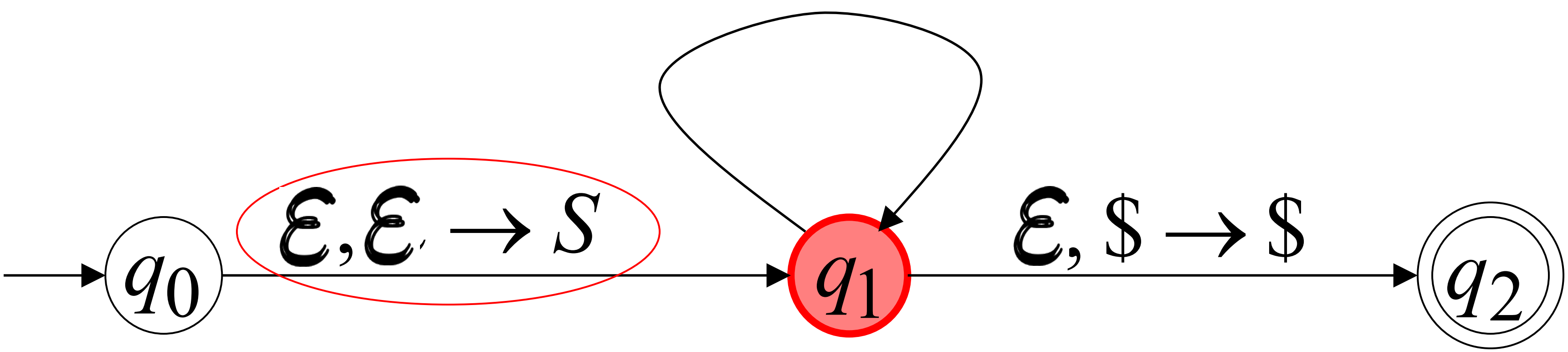
$\varepsilon, S \to b$

$\varepsilon, T \to Ta$ $\qquad a, a \to \varepsilon$

$\varepsilon, T \to \varepsilon$ $\qquad b, b \to \varepsilon$

Stack

S
$\$$



Derivation: $S \Rightarrow aSTb$

Input

a	b	a	b

Time 2

$\varepsilon, S \rightarrow aSTb$

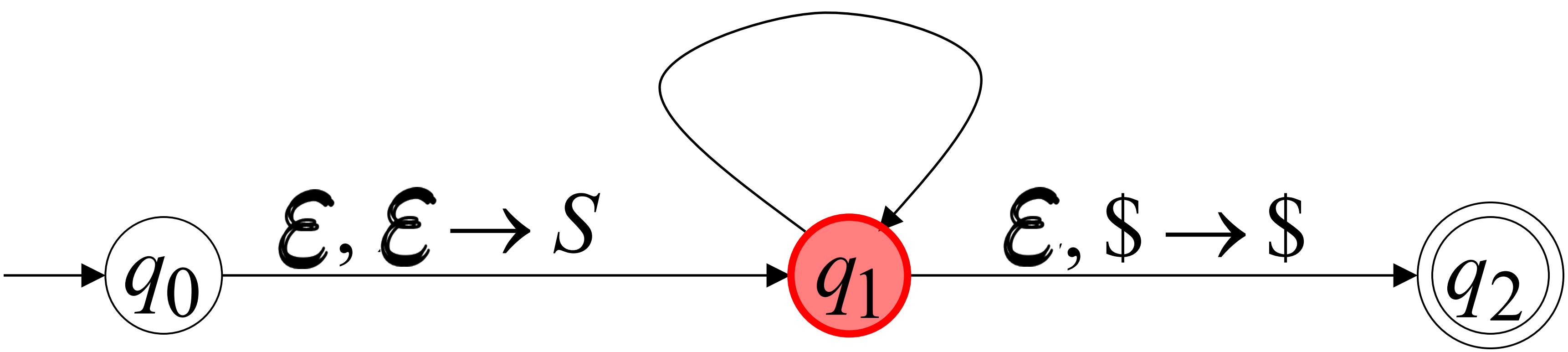
$\varepsilon, S \rightarrow b$

$\varepsilon, T \rightarrow Ta$ $\quad a, a \rightarrow \varepsilon$

$\varepsilon, T \rightarrow \varepsilon$ $\quad b, b \rightarrow \varepsilon$

a
S
T
b
$\$$

Stack



Derivation: $S \Rightarrow aSTb$

Input

a	b	a	b

Time 3

$\varepsilon, S \rightarrow aSTb$

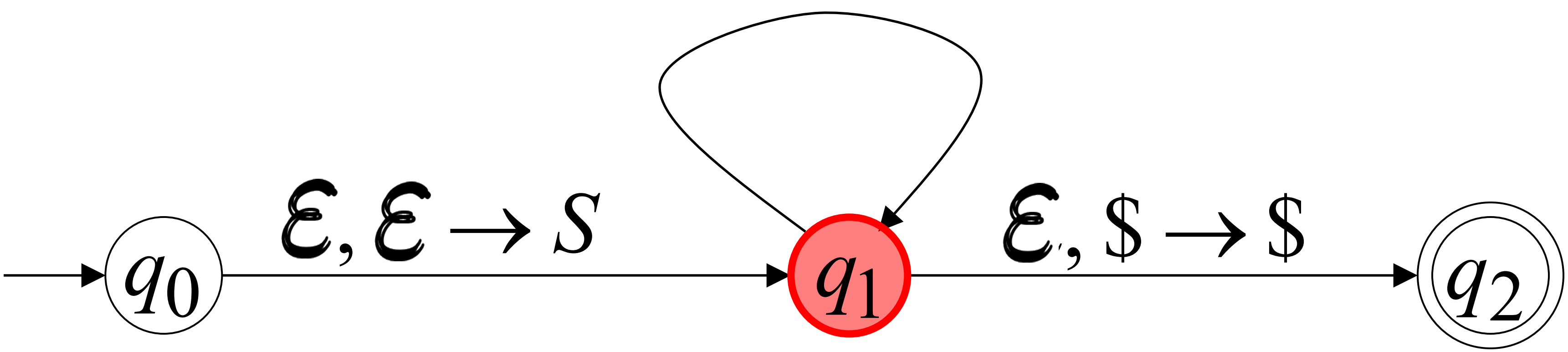
$\varepsilon, S \rightarrow b$

$\varepsilon, T \rightarrow Ta$ $\qquad a, a \rightarrow \varepsilon$

$\varepsilon, T \rightarrow \varepsilon$ $\qquad b, b \rightarrow \varepsilon$

Stack

$\cancel{a}$
S
T
b
$\$$



Derivation: $S \Rightarrow aSTb \Rightarrow abTb$

Input

a	b	a	b

Time 4

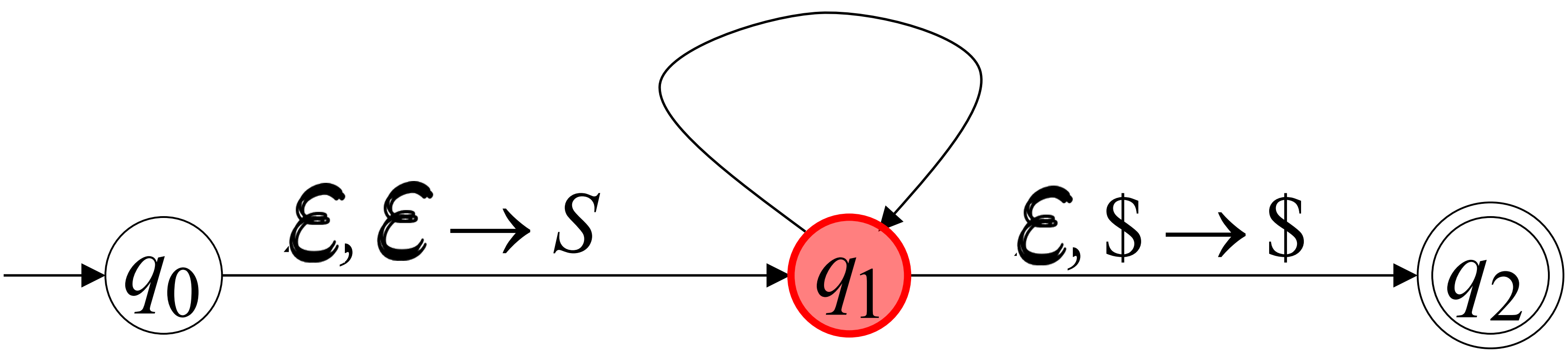
$\varepsilon, S \to aSTb$

$\varepsilon, S \to b$

$\varepsilon, T \to Ta$ $\quad a, a \to \varepsilon$

$\varepsilon, T \to \varepsilon$ $\quad b, b \to \varepsilon$

Stack
b
T
b
$\$$



Derivation: $S \Rightarrow aSTb \Rightarrow abTb$

Input

a	b	a	b

Stack

$\cancel{b}$
T
b
$\$$

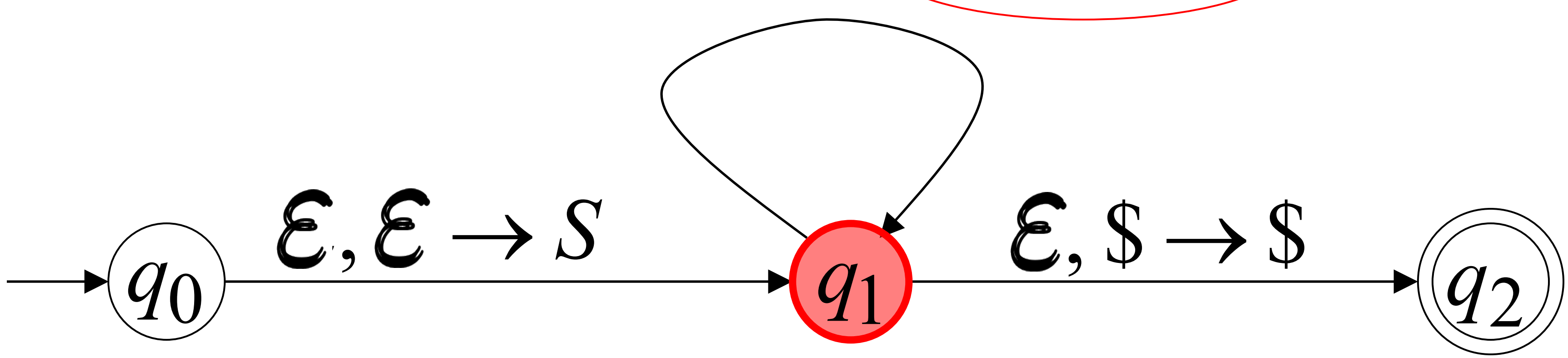
Time 5

$\varepsilon, S \rightarrow aSTb$

$\varepsilon, S \rightarrow b$

$\varepsilon, T \rightarrow Ta$ $\quad a, a \rightarrow \varepsilon$

$\varepsilon, T \rightarrow \varepsilon$ $\quad b, b \rightarrow \varepsilon$



Derivation: $S \Rightarrow aSTb \Rightarrow abTb \Rightarrow abTab$

Input

a	b	a	b

Time 6

$$\varepsilon, S \rightarrow aSTb$$

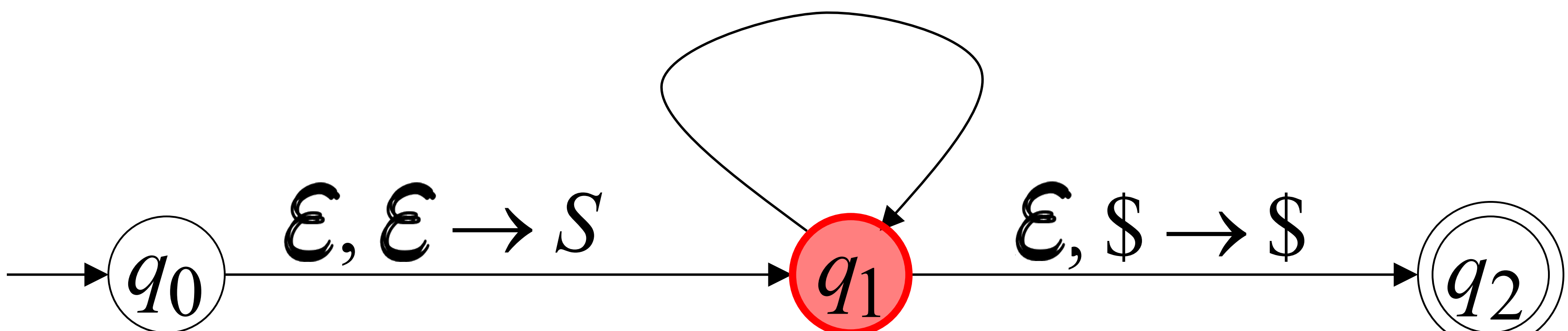
$$\varepsilon, S \rightarrow b$$

$$\varepsilon, T \rightarrow Ta \qquad a, a \rightarrow \varepsilon$$

$$\varepsilon, T \rightarrow \varepsilon \qquad b, b \rightarrow \varepsilon$$

T
a
b
$\$$

Stack



Derivation: $S \Rightarrow aSTb \Rightarrow abTb \Rightarrow abTab \Rightarrow abab$

Input

a	b	a	b

Stack

$\cancel{T}$
a
b
$\$$

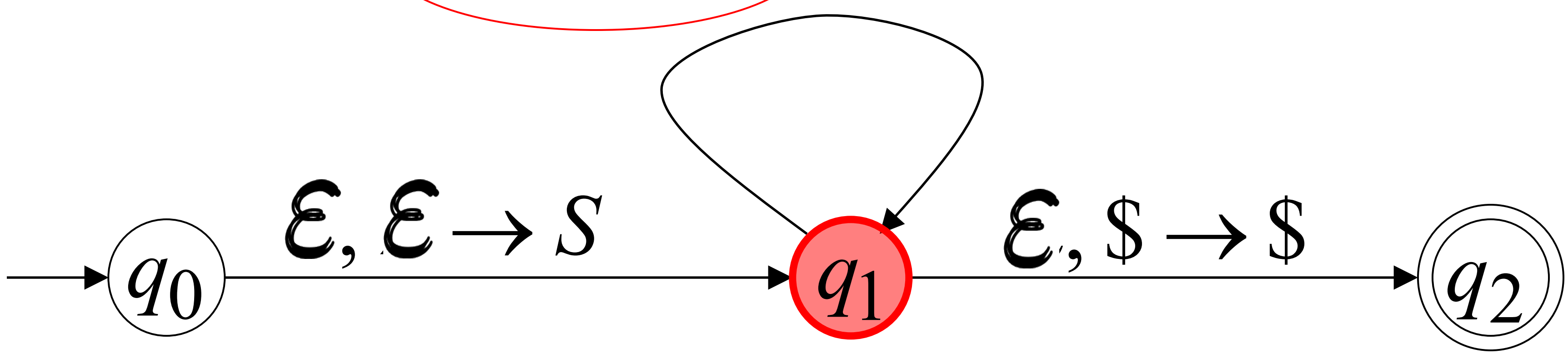
Time 7

$\varepsilon, S \to aSTb$

$\varepsilon, S \to b$

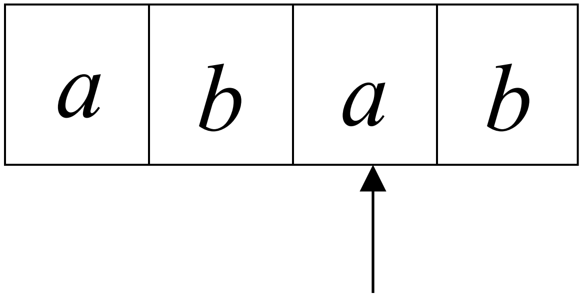
$\varepsilon, T \to Ta$ $\quad a, a \to \varepsilon$

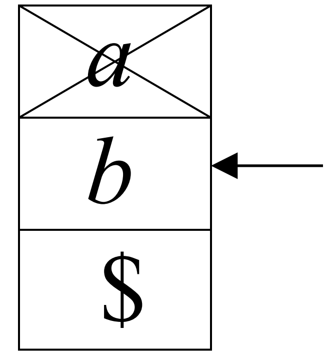
$\varepsilon, T \to \varepsilon$ $\quad b, b \to \varepsilon$



Derivation: $S \Rightarrow aSTb \Rightarrow abTb \Rightarrow abTab \Rightarrow abab$

Input





Stack

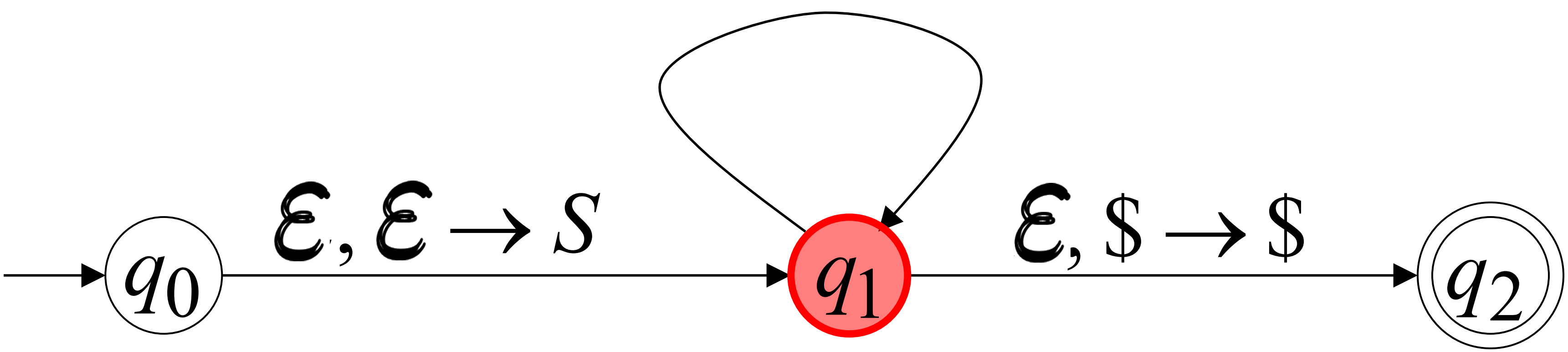
Time 8

$\varepsilon, S \rightarrow aSTb$

$\varepsilon, S \rightarrow b$

$\varepsilon, T \rightarrow Ta$ $\qquad a, a \rightarrow \varepsilon$

$\varepsilon, T \rightarrow \varepsilon$ $\qquad b, b \rightarrow \varepsilon$



Derivation: $S \Rightarrow aSTb \Rightarrow abTb \Rightarrow abTab \Rightarrow abab$

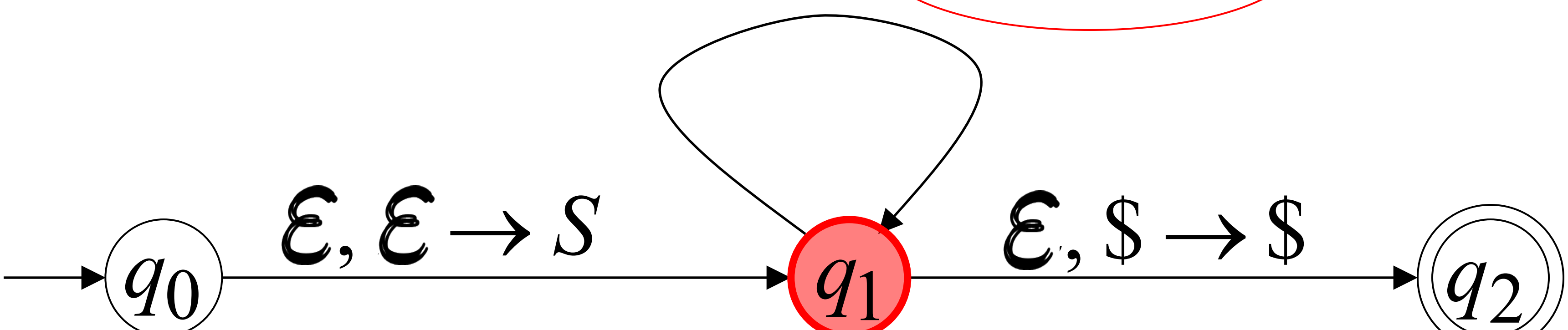
Input

a	b	a	b

Stack

$\cancel{b}$
$\$$

Time 9



Derivation: $S \Rightarrow aSTb \Rightarrow abTb \Rightarrow abTab \Rightarrow abab$

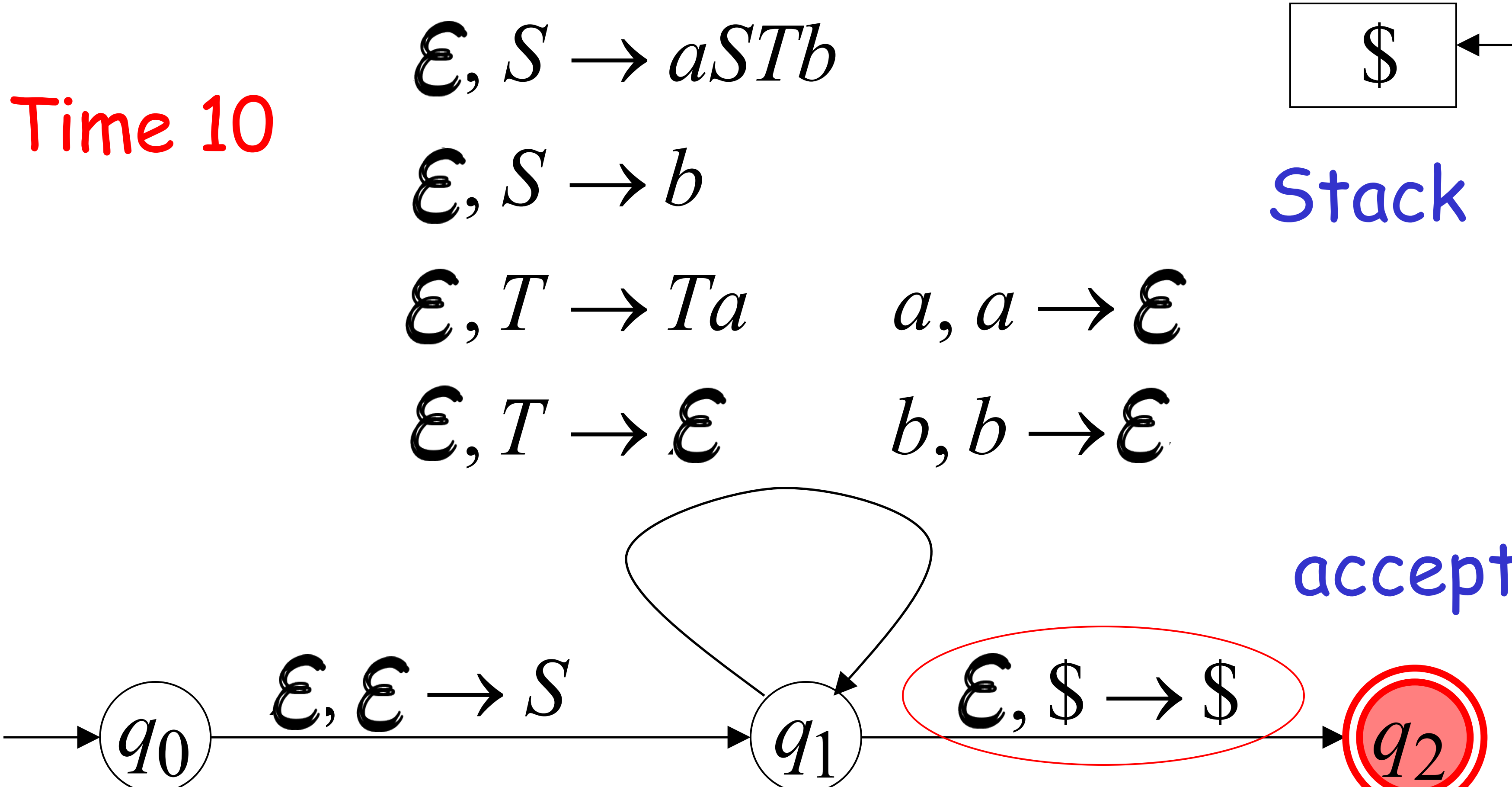
Input

a	b	a	b

Time 10

Stack

$\$$



Grammar Leftmost Derivation

PDA Computation

$$S$$

$$(q_0, abab, \$)$$
$$\succ (q_1, abab, S\$)$$

$$\Rightarrow aSTb$$

$$\succ (q_1, bab, STb\$)$$

$$\Rightarrow abTb$$

$$\succ (q_1, bab, bTb\$)$$
$$\succ (q_1, ab, Tb\$)$$

$$\Rightarrow abTab$$

$$\succ (q_1, ab, Tab\$)$$

$$\Rightarrow abab$$

$$\succ (q_1, ab, ab\$)$$
$$\succ (q_1, b, b\$)$$
$$\succ (q_1, \lambda, \$)$$
$$\succ (q_2, \lambda, \$)$$

In general, it can be shown that:

Grammar G
generates
string w

$$S \overset{*}{\Rightarrow} w$$

If and Only if

PDA M
accepts w

$$(q_0, w, \$) \succ (q_2, \varepsilon, \$)$$

Therefore $L(G) = L(M)$

Proof - step 2

Convert

PDAs
to
Context-Free Grammars

Take an arbitrary PDA M

We will convert M
to a context-free grammar G such that:

$$L(M) = L(G)$$

First modify PDA M so that:

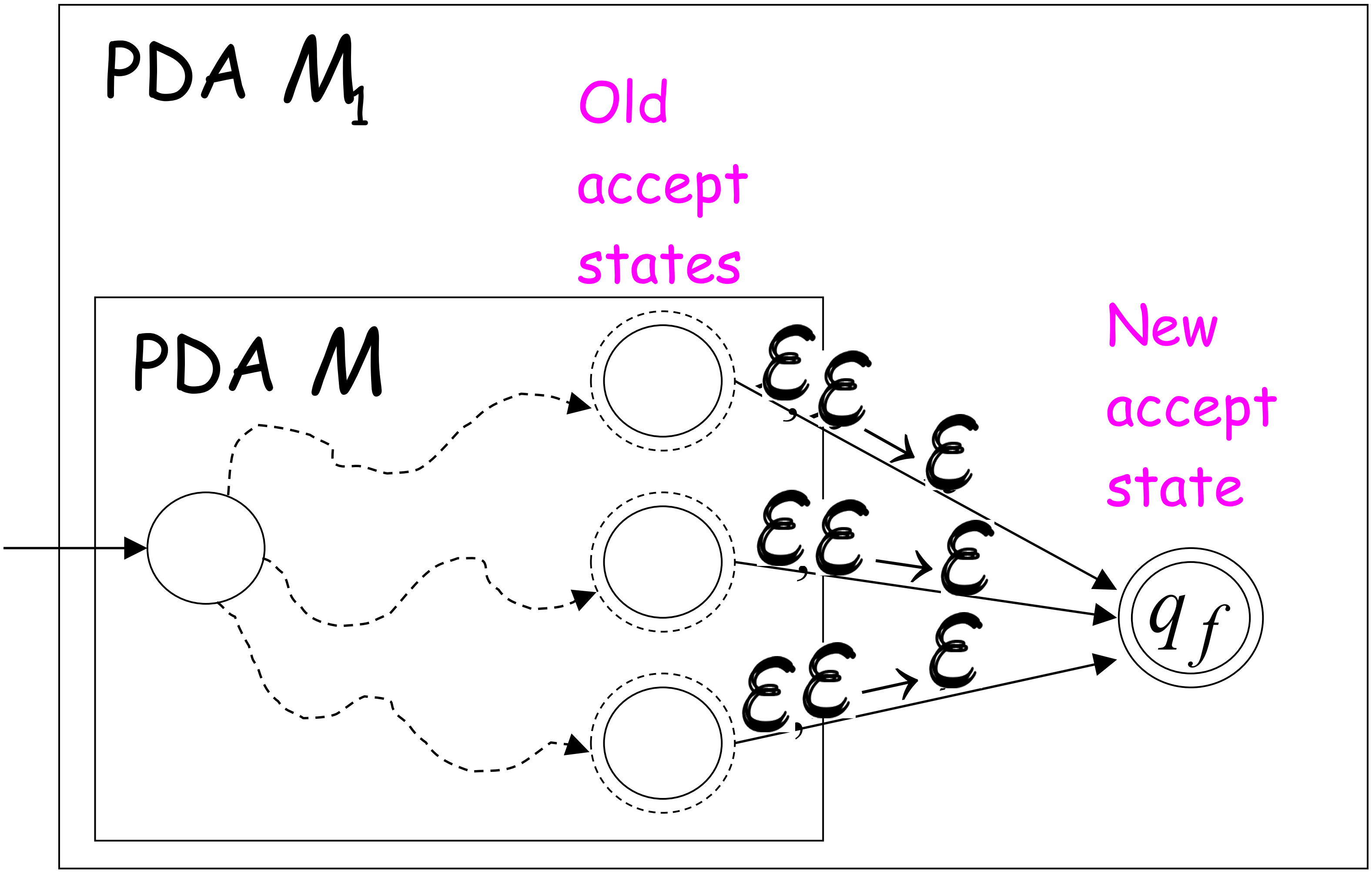
1. The PDA has a single accept state

2. Use new initial stack symbol #

3. On acceptance the stack contains only stack symbol #

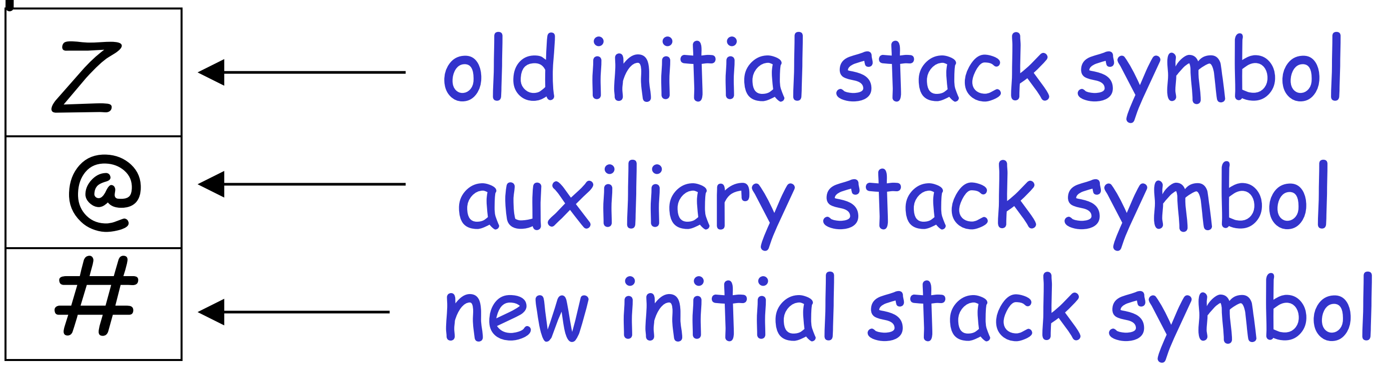
4. Each transition either pushes a symbol or pops a symbol but not both together

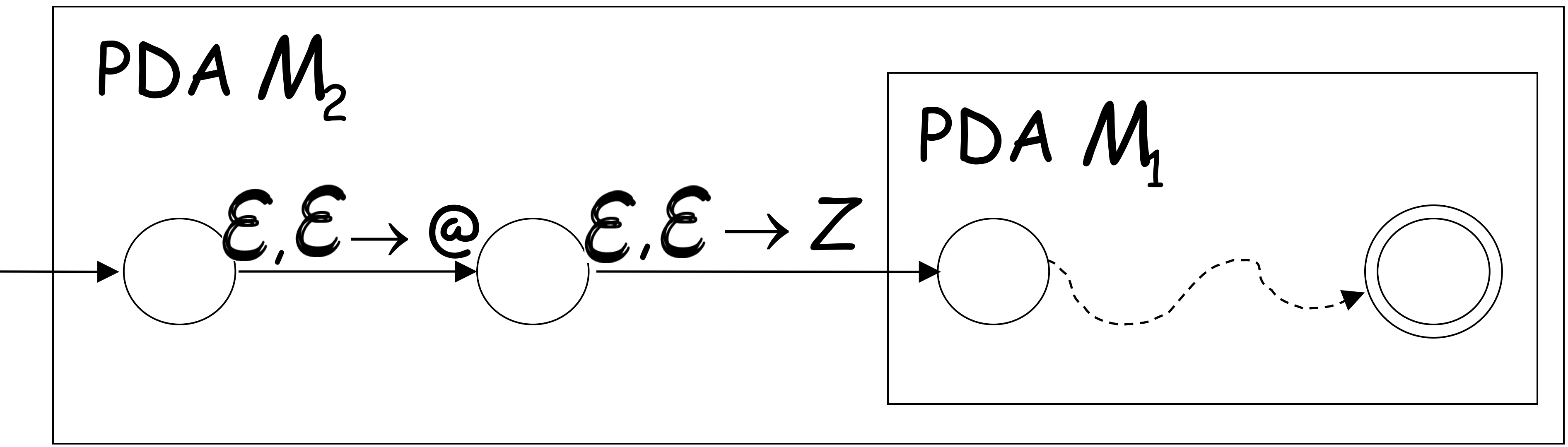
1. The PDA has a single accept state



2. Use new initial stack symbol

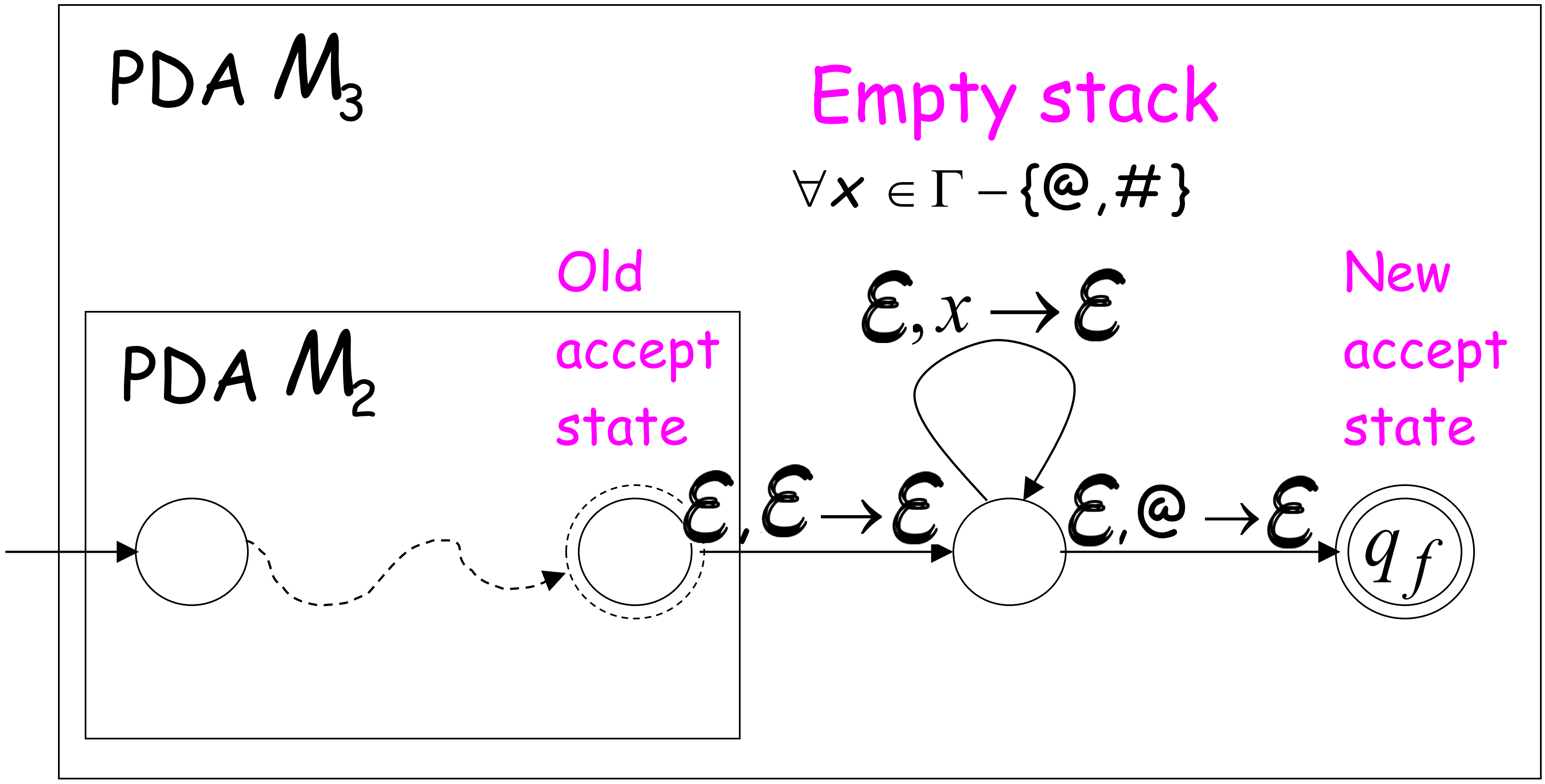
Top of stack





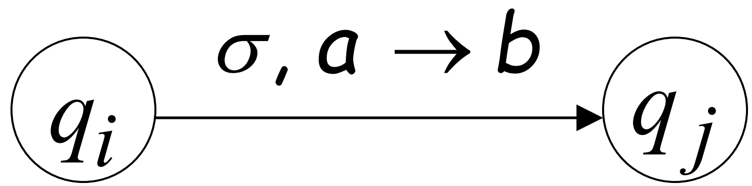
M_1 still thinks that Z is the initial stack

3. On acceptance the stack contains only stack symbol

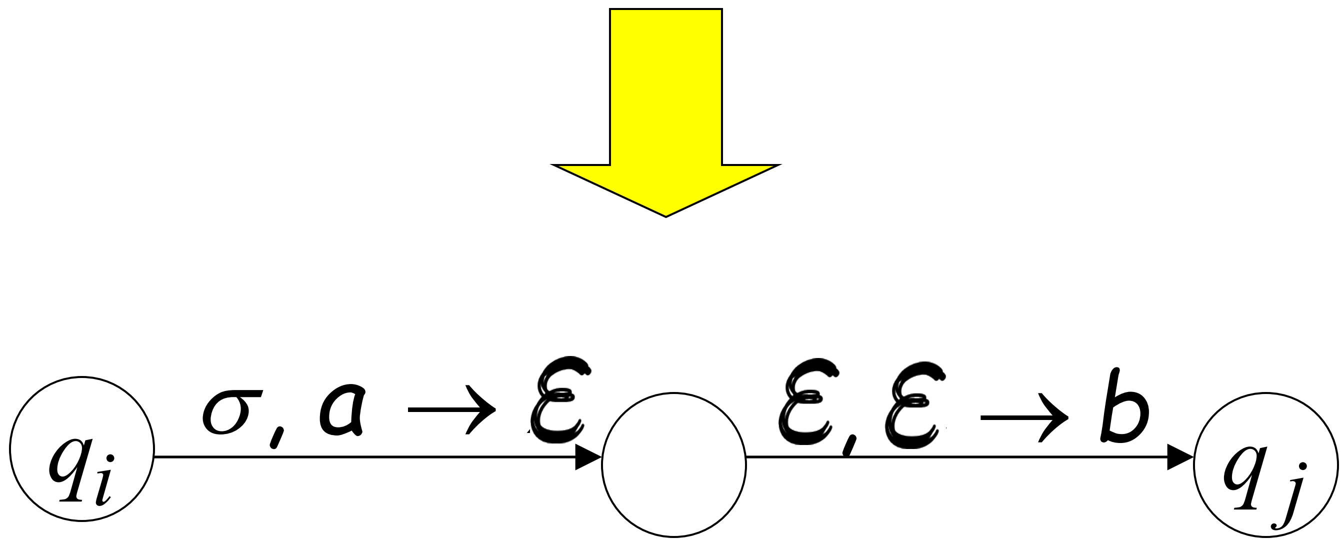


4. Each transition either pushes a symbol or pops a symbol but not both together

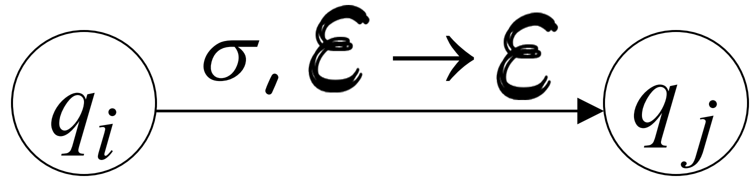
PDA M_3



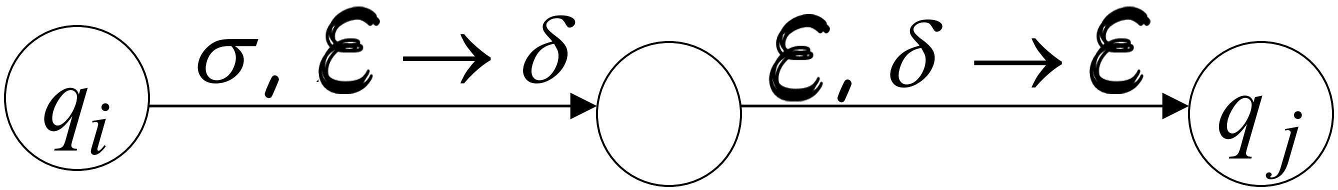
PDA M_4



PDA M_3



PDA M_4

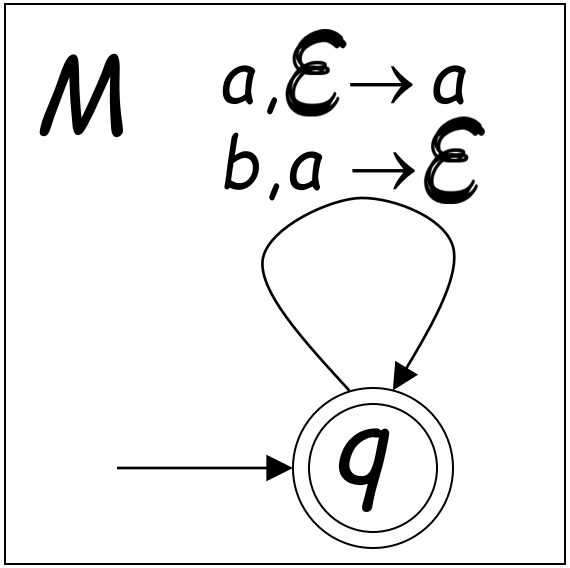


Where δ is a symbol of the stack alphabet

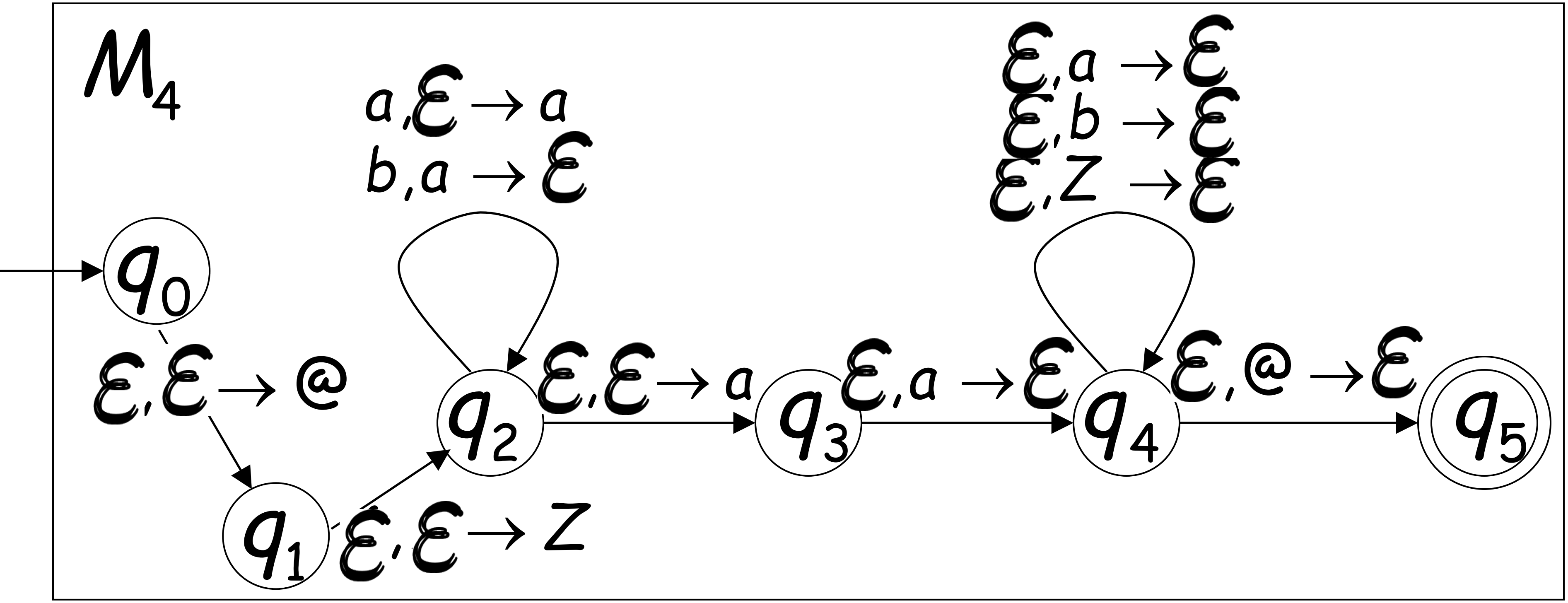
PDA M_4 is the final modified PDA

Note that the new initial stack symbol #
is never used in any transition

Example:



M
$a, \varepsilon \rightarrow a$
$b, a \rightarrow \varepsilon$
q



M_4
q_0
$\varepsilon, \varepsilon \rightarrow @$
q_1
$\varepsilon, \varepsilon \rightarrow Z$
$a, \varepsilon \rightarrow a$
$b, a \rightarrow \varepsilon$
q_2
$\varepsilon, \varepsilon \rightarrow a$
q_3
$\varepsilon, a \rightarrow \varepsilon$
$\varepsilon, a \rightarrow \varepsilon$
$\varepsilon, b \rightarrow \varepsilon$
$\varepsilon, Z \rightarrow \varepsilon$
q_4
$\varepsilon, @ \rightarrow \varepsilon$
q_5

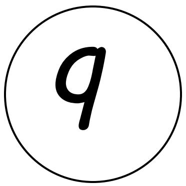
Grammar Construction

Variables: A_{q_i, q_j}

States of PDA

PDA

Kind 1: for each state

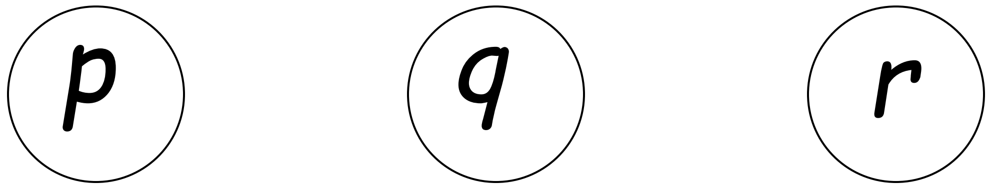


Grammar

$$A_{qq} \to \varepsilon,$$

PDA

Kind 2: for every three states



Grammar

$$A_{pq} \rightarrow A_{pr} A_{rq}$$

PDA

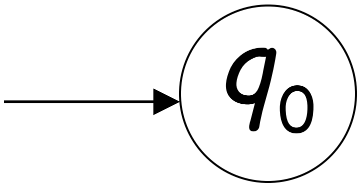
Kind 3: for every pair of such transitions

Grammar

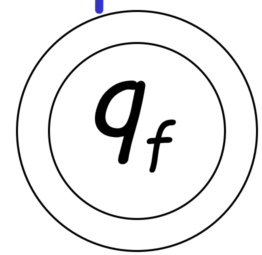
$$A_{pq} \rightarrow a A_{rs} b$$

PDA

Initial state



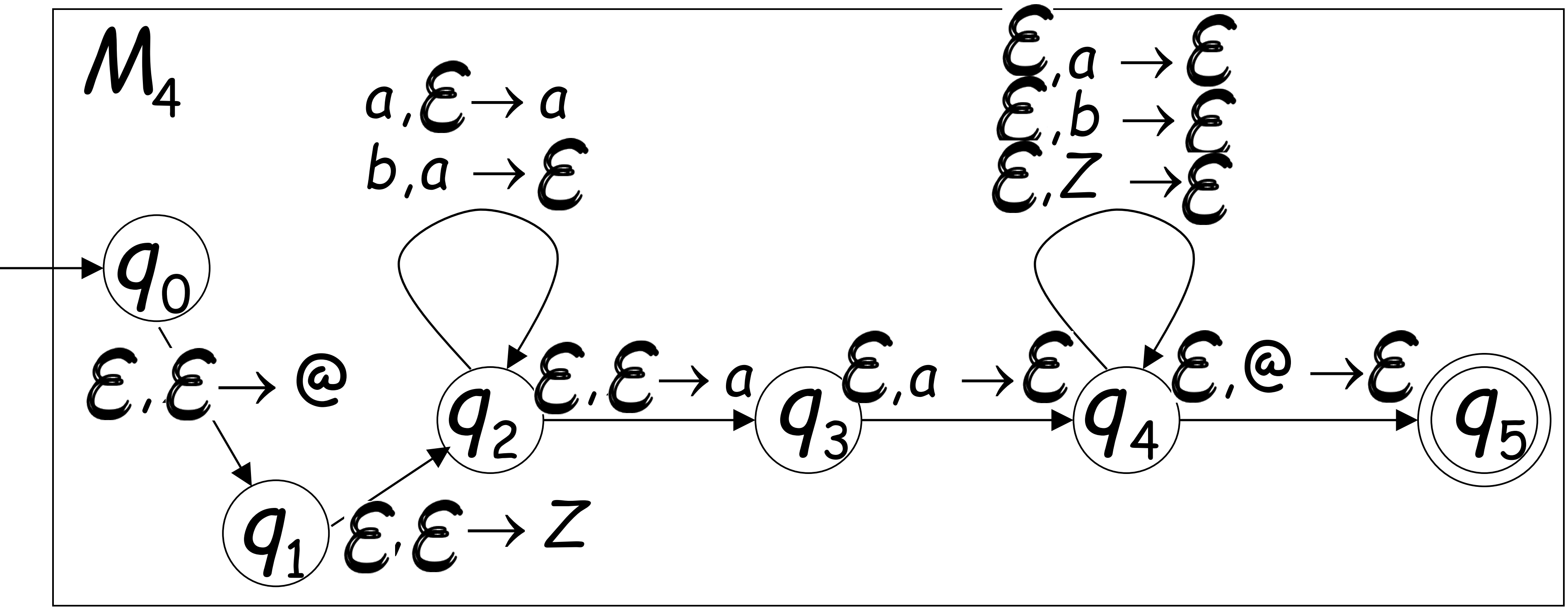
Accept state



Grammar

Start variable $A_{q_0 q_f}$

Example: PDA



M_4
q_0
q_1
q_2
q_3
q_4
q_5
$\varepsilon, \varepsilon \rightarrow @$
$\varepsilon, \varepsilon \rightarrow Z$
$a, \varepsilon \rightarrow a$
$b, a \rightarrow \varepsilon$
$\varepsilon, \varepsilon \rightarrow a$
$\varepsilon, a \rightarrow \varepsilon$
$\varepsilon, a \rightarrow \varepsilon$
$\varepsilon, b \rightarrow \varepsilon$
$\varepsilon, Z \rightarrow \varepsilon$
$\varepsilon, @ \rightarrow \varepsilon$

Grammar

Kind 1: from single states

$$A_{q_0 q_0} \rightarrow \varepsilon$$

$$A_{q_1 q_1} \rightarrow \varepsilon$$

$$A_{q_2 q_2} \rightarrow \varepsilon$$

$$A_{q_3 q_3} \rightarrow \varepsilon$$

$$A_{q_4 q_4} \rightarrow \varepsilon$$

$$A_{q_5 q_5} \rightarrow \varepsilon$$

Kind 2: from triplets of states

$$A_{q_0q_0} \rightarrow A_{q_0q_0}A_{q_0q_0} \mid A_{q_0q_1}A_{q_1q_0} \mid A_{q_0q_2}A_{q_2q_0} \mid A_{q_0q_3}A_{q_3q_0} \mid A_{q_0q_4}A_{q_4q_0} \mid A_{q_0q_5}A_{q_5q_0}$$

$$A_{q_0q_1} \rightarrow A_{q_0q_0}A_{q_0q_1} \mid A_{q_0q_1}A_{q_1q_1} \mid A_{q_0q_2}A_{q_2q_1} \mid A_{q_0q_3}A_{q_3q_1} \mid A_{q_0q_4}A_{q_4q_1} \mid A_{q_0q_5}A_{q_5q_1}$$

$\vdots$

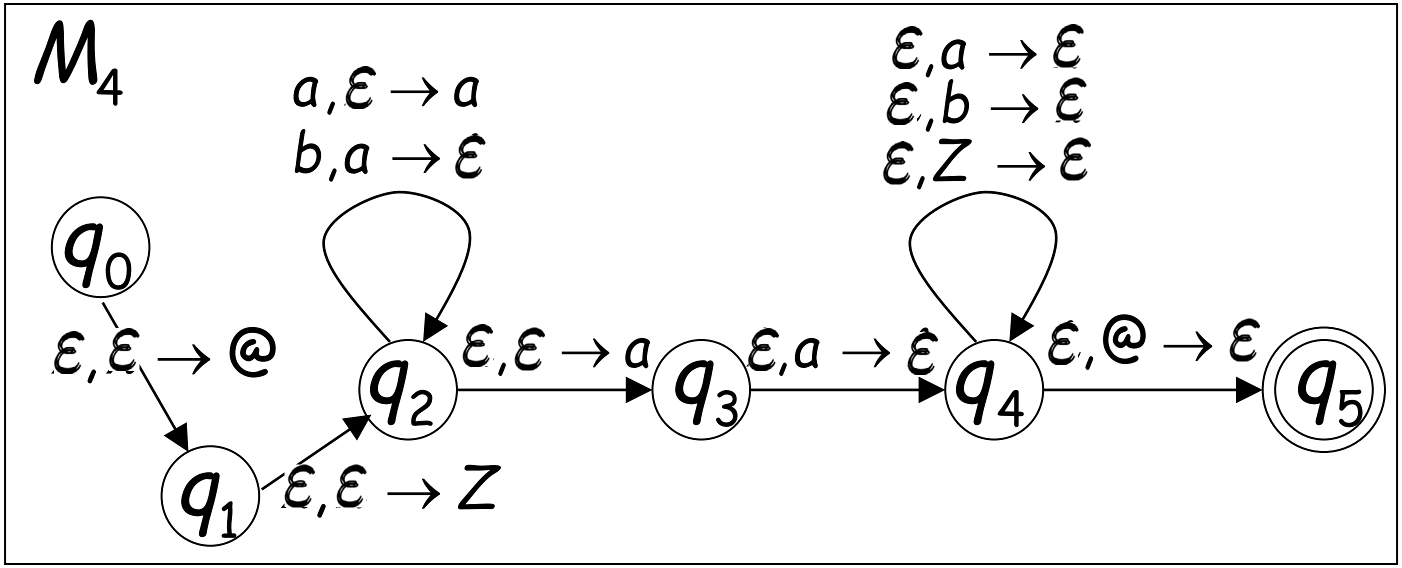
$$A_{q_0q_5} \rightarrow A_{q_0q_0}A_{q_0q_5} \mid A_{q_0q_1}A_{q_1q_5} \mid A_{q_0q_2}A_{q_2q_5} \mid A_{q_0q_3}A_{q_3q_5} \mid A_{q_0q_4}A_{q_4q_5} \mid A_{q_0q_5}A_{q_5q_5}$$

$\vdots$

$$A_{q_5q_5} \rightarrow A_{q_5q_0}A_{q_0q_5} \mid A_{q_5q_1}A_{q_1q_5} \mid A_{q_5q_2}A_{q_2q_5} \mid A_{q_5q_3}A_{q_3q_5} \mid A_{q_5q_4}A_{q_4q_5} \mid A_{q_5q_5}A_{q_5q_5}$$

Start variable $A_{q_0q_5}$

Kind 3: from pairs of transitions



$$A_{q_0q_5} \to A_{q_1q_4}$$

$$A_{q_1q_4} \to A_{q_2q_4}$$

$$A_{q_2q_4} \to aA_{q_2q_4}$$

$$A_{q_2q_2} \to aA_{q_2q_2}b$$

$$A_{q_2q_4} \to aA_{q_2q_3}$$

$$A_{q_2q_2} \to A_{q_3q_2}b$$

$$A_{q_2q_4} \to A_{q_3q_3}$$

$$A_{q_2q_4} \to A_{q_3q_4}$$

Suppose that a PDA M is converted to a context-free grammar G

We need to prove that $L(G) = L(M)$

or equivalently

$L(G) \subseteq L(M)$ $\qquad$ $L(G) \supseteq L(M)$

$L(G) \subseteq L(M)$

We need to show that if G has derivation:

$$A_{q_0 q_f} \stackrel{*}{\Rightarrow} w \qquad \text{(string of terminals)}$$

Then there is an accepting computation in M :

$$(q_0, w, \#) \stackrel{*}{\succ} (q_f, \varepsilon, \#)$$

with input string w

We will actually show that if G has derivation:

$$A_{pq} \overset{*}{\Rightarrow} w$$

Then there is a computation in M :

$$(p, w, \varepsilon) \overset{*}{\succ} (q, \varepsilon, \varepsilon)$$

Therefore:

$$A_{q_0 q_f} \overset{*}{\Rightarrow} w$$

$$(q_0, w, \varepsilon) \overset{*}{\succ} (q_f, \varepsilon, \varepsilon)$$

Since there is no transition with the # symbol

$$(q_0, w, \#) \overset{*}{\succ} (q_f, \varepsilon, \#)$$

Lemma:

If $A_{pq} \overset{*}{\Rightarrow} w$ (string of terminals)

then there is a computation from state p to state q on string w which leaves the stack empty:

$$(p, w, \varepsilon) \overset{*}{\succ} (q, \varepsilon, \varepsilon)$$

Proof Intuition:

$$A_{pq} \Rightarrow \cdots \Rightarrow w$$

Type 2

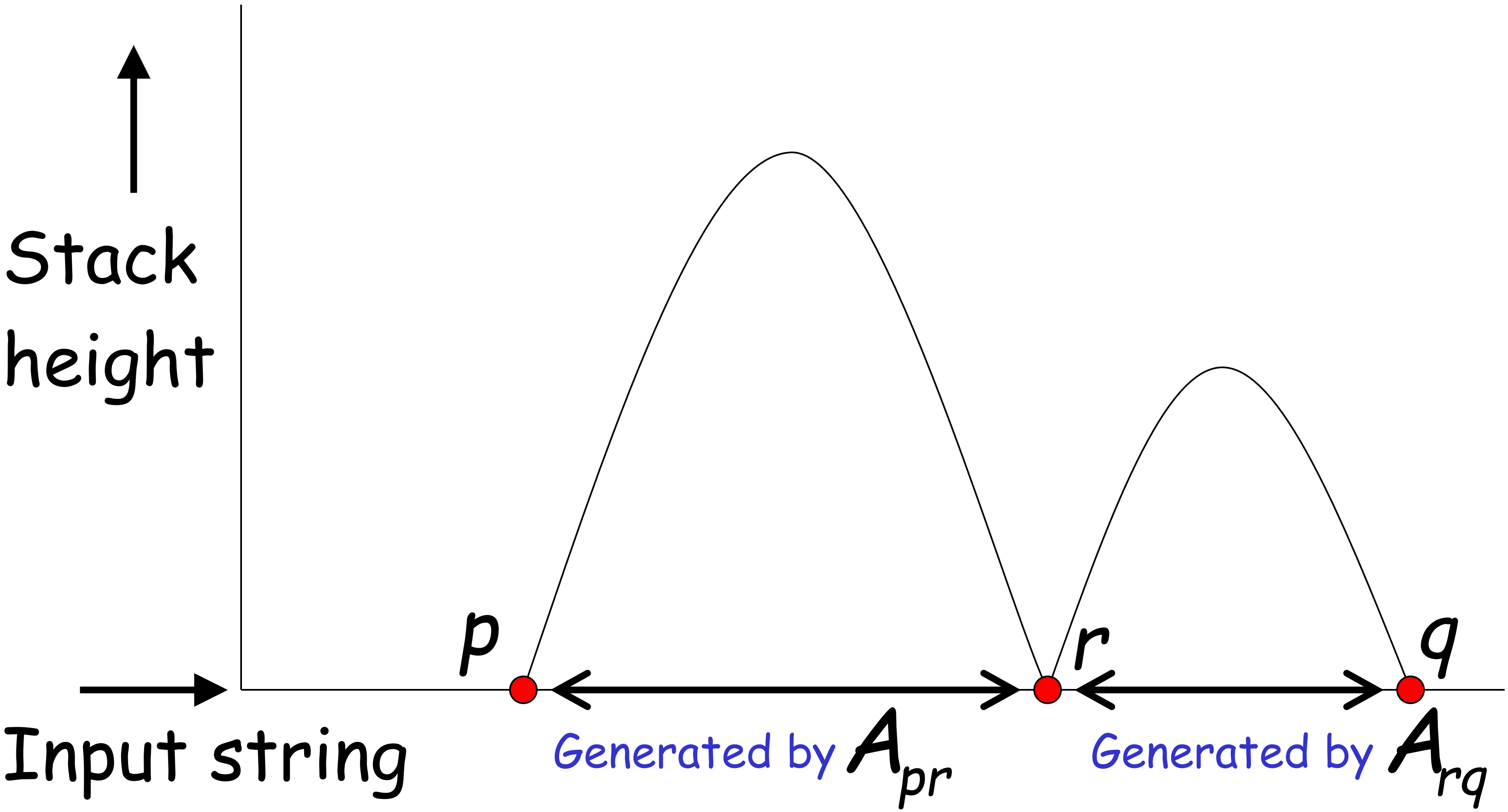
Case 1: $$A_{pq} \Rightarrow A_{pr}A_{rq} \Rightarrow \cdots \Rightarrow w$$

Type 3

Case 2: $$A_{pq} \Rightarrow aA_{rs}b \Rightarrow \cdots \Rightarrow w$$

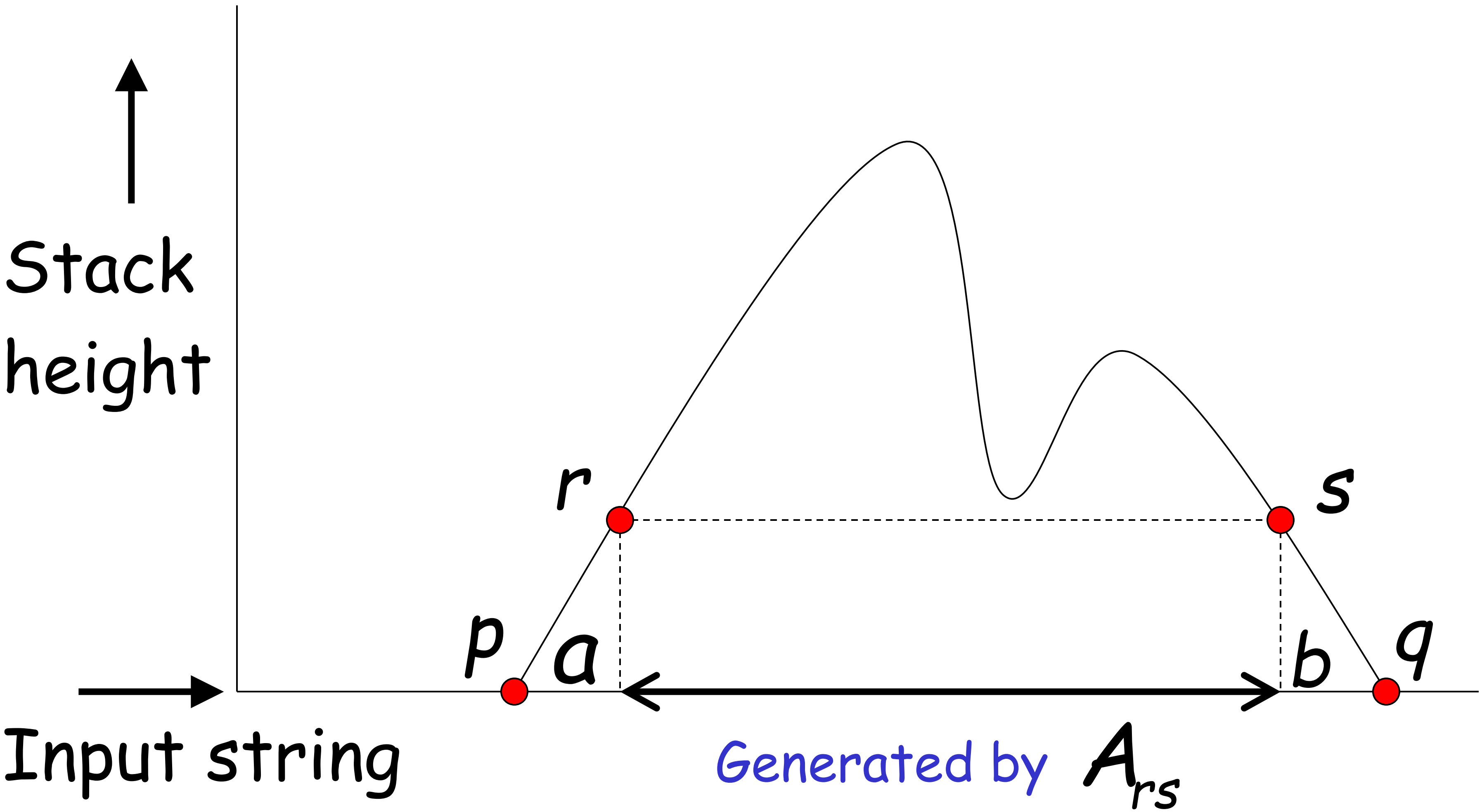
Type 2

Case 1: $A_{pq} \Rightarrow A_{pr} A_{rq} \Rightarrow \cdots \Rightarrow w$



Type 3

Case 2: $A_{pq} \Rightarrow aA_{rs}b \Rightarrow \cdots \Rightarrow w$



Formal Proof:

We formally prove this claim by induction on the number of steps in derivation:

$$A_{pq} \underbrace{\Rightarrow \cdots \Rightarrow}_{\text{number of steps}} w$$

Induction Basis: $A_{pq} \Rightarrow w$

(one derivation step)

A Kind 1 production must have been used:

$$A_{pp} \to \varepsilon$$

Therefore, $p = q$ and $w = \varepsilon$

This computation of PDA trivially exists:

$$(p,\varepsilon,\varepsilon) \overset{*}{\succ} (p,\varepsilon,\varepsilon)$$

Induction Hypothesis:

$$\underbrace{A_{pq} \Rightarrow \cdots \Rightarrow}_{k} w$$

k derivation steps

suppose it holds:

$$(p, w, \varepsilon) \overset{*}{\succ} (q, \varepsilon, \varepsilon)$$

Induction Step:

$$\underbrace{A_{pq} \Rightarrow \cdots \Rightarrow w}_{k+1 \text{ derivation steps}}$$

We have to show:

$$(p, w, \varepsilon) \overset{*}{\succ} (q, \varepsilon, \varepsilon)$$

$$\underbrace{A_{pq} \Rightarrow \cdots \Rightarrow}_{k+1} w$$

$k+1$ derivation steps

Type 2

Case 1: $A_{pq} \Rightarrow A_{pr}A_{rq} \Rightarrow \cdots \Rightarrow w$

Type 3

Case 2: $A_{pq} \Rightarrow aA_{rs}b \Rightarrow \cdots \Rightarrow w$

Type 2

Case 1: $A_{pq} \underbrace{\Rightarrow A_{pr} A_{rq} \Rightarrow \cdots \Rightarrow}_{k+1 \text{ steps}} w$

We can write $w = yz$

$A_{pr} \underbrace{\Rightarrow \cdots \Rightarrow}_{\text{At most } k \text{ steps}} y$

$A_{rq} \underbrace{\Rightarrow \cdots \Rightarrow}_{\text{At most } k \text{ steps}} z$

$$\underbrace{A_{pr} \Rightarrow \cdots \Rightarrow}_{\text{At most } k \text{ steps}} y$$

From induction hypothesis, in PDA:

$$(p, y, \varepsilon) \overset{*}{\succ} (r, \varepsilon, \varepsilon)$$

$$\underbrace{A_{rq} \Rightarrow \cdots \Rightarrow}_{\text{At most } k \text{ steps}} z$$

From induction hypothesis, in PDA:

$$(r, z, \varepsilon) \overset{*}{\succ} (q, \varepsilon, \varepsilon)$$

$$(p, y, \varepsilon) \overset{*}{\succ} (r, \varepsilon, \varepsilon) \qquad (r, z, \varepsilon) \overset{*}{\succ} (q, \varepsilon, \varepsilon)$$

$$(p, yz, \varepsilon) \overset{*}{\succ} (r, z, \varepsilon) \overset{*}{\succ} (q, \varepsilon, \varepsilon)$$

since $w = yz$

$$(p, w, \varepsilon) \overset{*}{\succ} (q, \varepsilon, \varepsilon)$$

Type 3

Case 2: $A_{pq} \underbrace{\Rightarrow aA_{rs}b \Rightarrow \cdots \Rightarrow}_{k+1 \text{ steps}} w$

We can write $w = ayb$

$A_{rs} \underbrace{\Rightarrow \cdots \Rightarrow}_{\text{At most } k \text{ steps}} y$

$$\underbrace{A_{rs} \Rightarrow \cdots \Rightarrow}_{\text{At most } k \text{ steps}} y$$

From induction hypothesis, the PDA has computation:

$$(r, y, \varepsilon) \overset{*}{\succ} (s, \varepsilon, \varepsilon)$$

Type 3

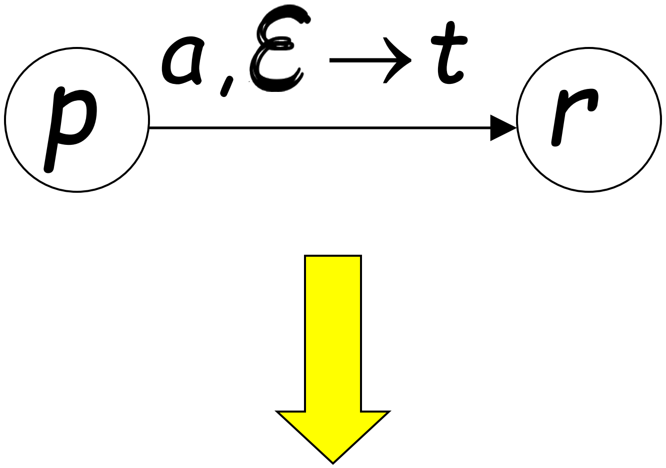
$$A_{pq} \Rightarrow aA_{rs}b \Rightarrow \cdots \Rightarrow w$$

Grammar contains production

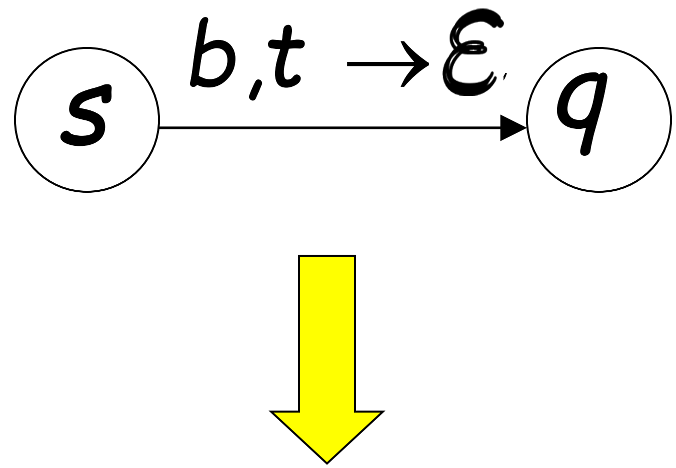
$$A_{pq} \rightarrow aA_{rs}b$$

And PDA Contains transitions

$$p \xrightarrow{a,\varepsilon \rightarrow t} r \qquad\qquad s \xrightarrow{b,t \rightarrow \varepsilon} q$$



$$(p, ayb, \varepsilon) \succ (r, yb, t)$$



$$(s, b, t) \succ (q, \varepsilon, \varepsilon)$$

We know

$$(r,y,\varepsilon) \overset{*}{\succ} (s,\varepsilon,\varepsilon) \quad \Longrightarrow \quad (r,yb,t) \overset{*}{\succ} (s,b,t)$$

$$(p,ayb,\varepsilon) \succ (r,yb,t)$$

We also know

$$(s,b,t) \succ (q,\varepsilon,\varepsilon)$$

Therefore:

$$(p,ayb,\varepsilon) \succ (r,yb,t) \overset{*}{\succ} (s,b,t) \succ (q,\varepsilon,\varepsilon)$$

$$(p, ayb, \varepsilon) \succ (r, yb, t) \overset{*}{\succ} (s, b, t) \succ (q, \varepsilon, \varepsilon)$$

since $w = ayb$

$$(p, w, \varepsilon) \overset{*}{\succ} (q, \varepsilon, \varepsilon)$$

END OF PROOF

So far we have shown:

$$L(G) \subseteq L(M)$$

With a similar proof we can show

$$L(G) \supseteq L(M)$$

Therefore: $L(G) = L(M)$