

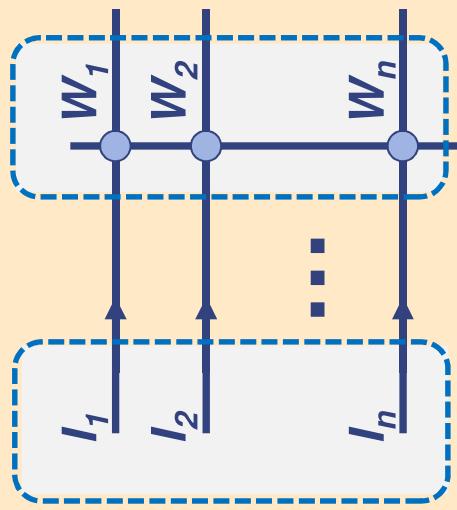
## 2.9 „Convolutional Networks“

- Implementierung eines Neuronalen Netzes
- Ersatz der allgemeinen Matrix-Multiplikation durch Faltungsoperationen in mindestens einer Schicht
  - Spärliche Verbindungen
    - „sparse interaction“, „sparse connectivity“, „sparse weights“
  - Gemeinsame Verwendung von Parametern
    - „parameter sharing“
  - Equivariante Representation von Information
- Poolbildung, Vereinigungsoperationen
  - „Pooling“



Ansatz für  $M$ : gewichtete Summation :

Input: Pixel  $\in R(x_0, y_0)$



Linien-Detektion am Ort  $(x_0, y_0)$

### Formulierung 2-dimensional

$$M = \sum_{(x,y) \in R(x_0, y_0)} W(x, y) \cdot I(x, y)$$

Linien-Detektion am Ort  $(x_1, y_1)$      $x_1 = x_0 + \Delta x$      $y_1 = y_0 + \Delta y$

... identische Methode

... Verschiebung Bild

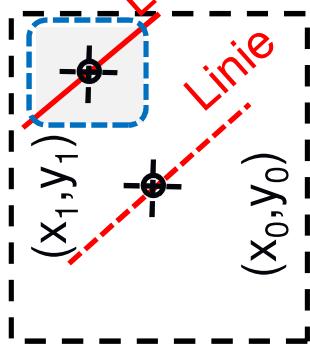
$$M = \sum_{(x,y) \in R(x_0, y_0)} W(x, y) \cdot I(x - \Delta x, y - \Delta y)$$

$$s = \sum_{k=1}^n W_k \cdot I_k$$

lin.

**1-d Prototyp**

$$M = s$$

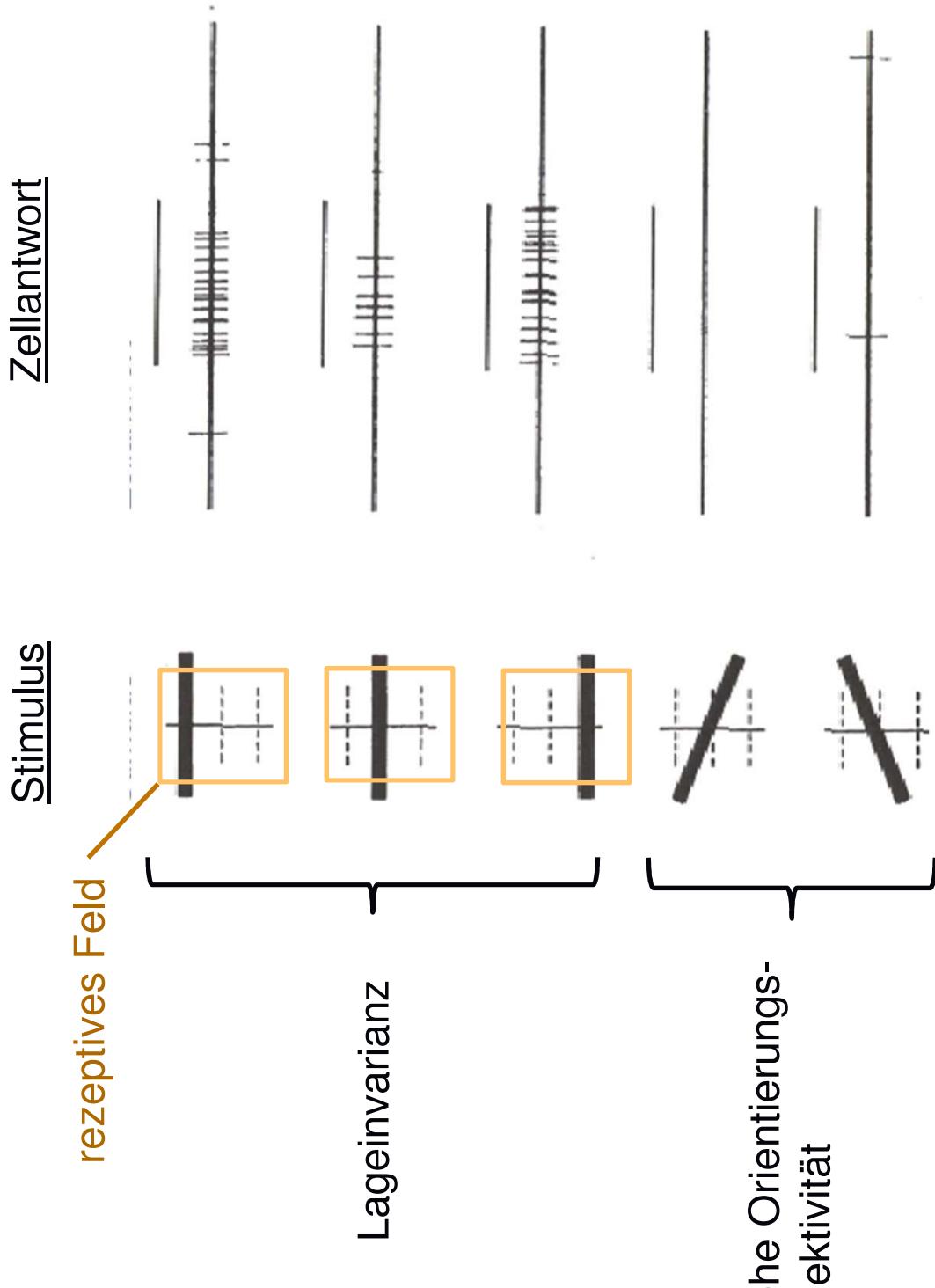


Faltung (Convolution): Filterkern wird an unterschiedlichen Positionen appliziert

$$M(\Delta x, \Delta y) = \sum_{(x,y) \in R(x_0, y_0)} W(x, y) \cdot I(x - \Delta x, y - \Delta y)$$

- Neben den „**Simple Cells**“, die ein nahezu lineare Übertragungsverhalten aufweisen, gibt es sog. „**Complex Cells**“, die ein ausgeprägt nicht-lineares Verhalten zeigen.

- Beispiel



## „Complex Cells“ finden sich u.a. im

V1 (primärer visueller Kortex)

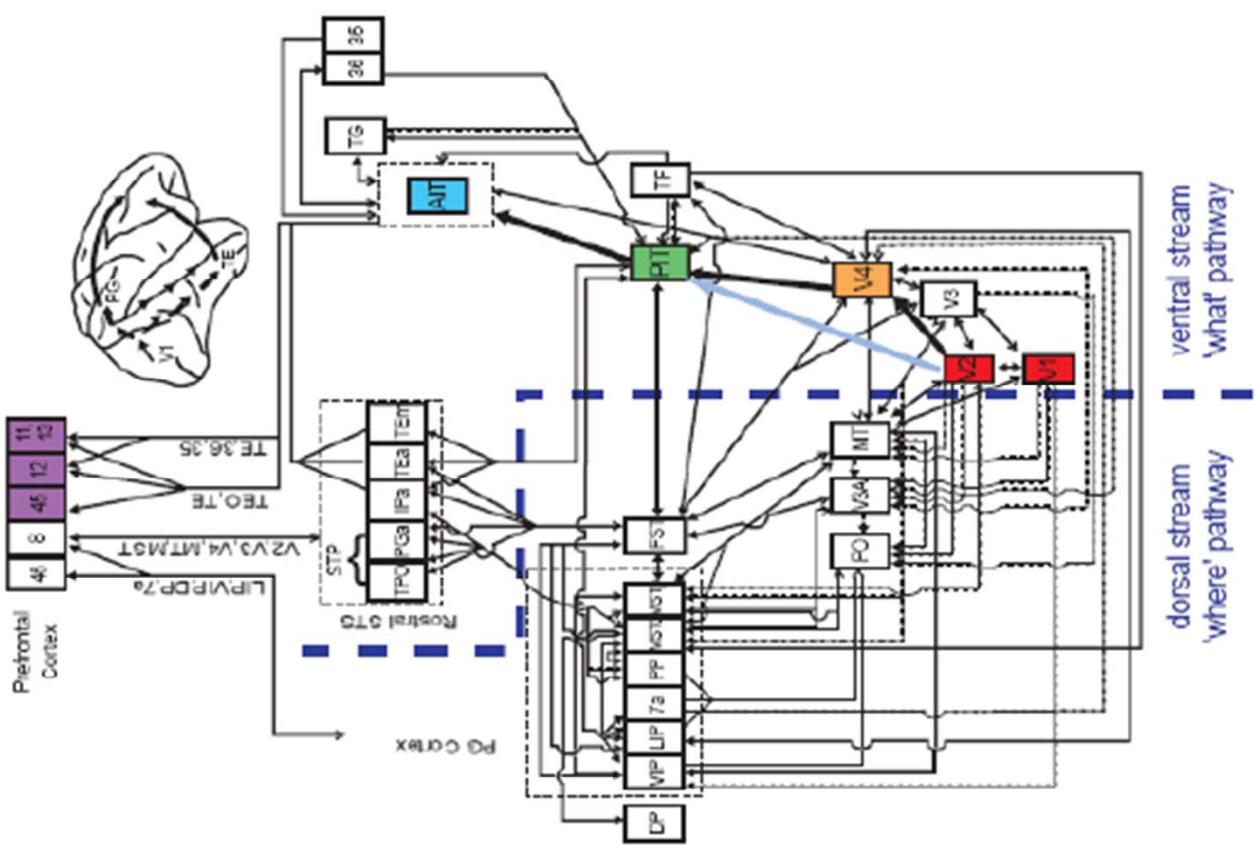
V2 (sekundärer visueller Kortex)

V3 (Brodmann Area 19)

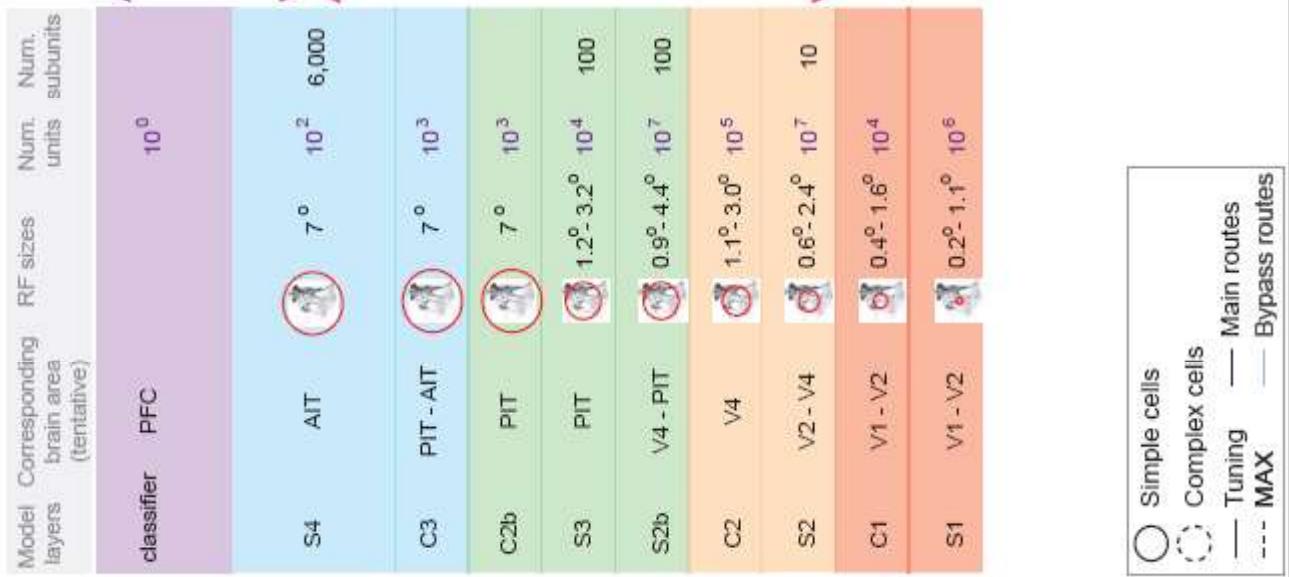
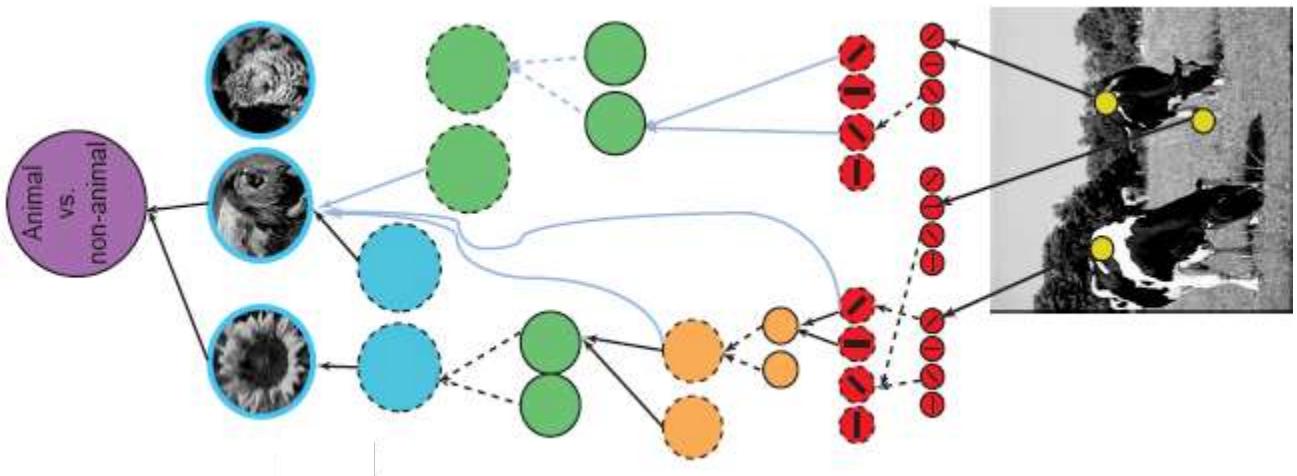
des visuellen Pfades.

### Charakteristik

- z.B. sensitiv für orientierte Linien
- Translationssymmetrie  
(Invarianz gegenüber kleinen räumlichen Verschiebungen des Stimulus)



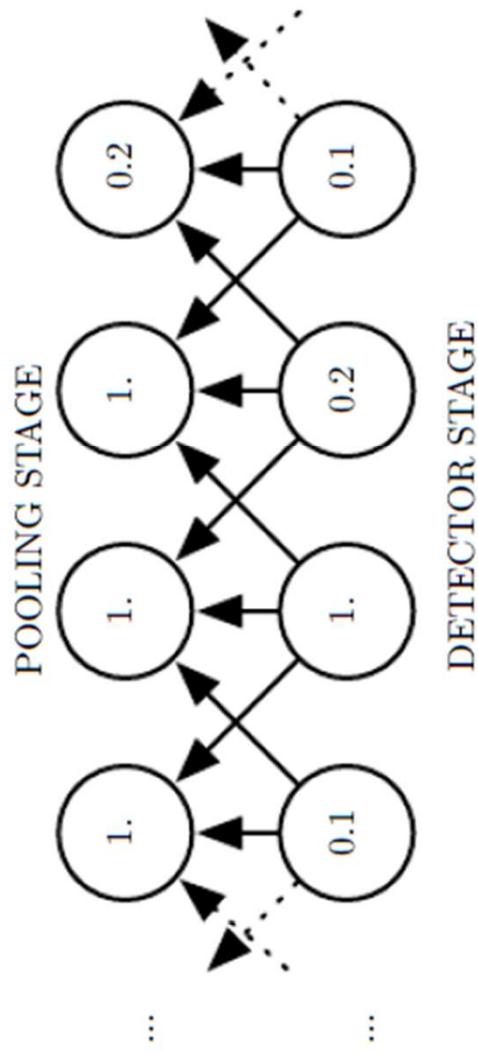
## Hierachie



**Ansatz: „Complex Cells“**  
können durch eine Pooling-  
operation modelliert werden.

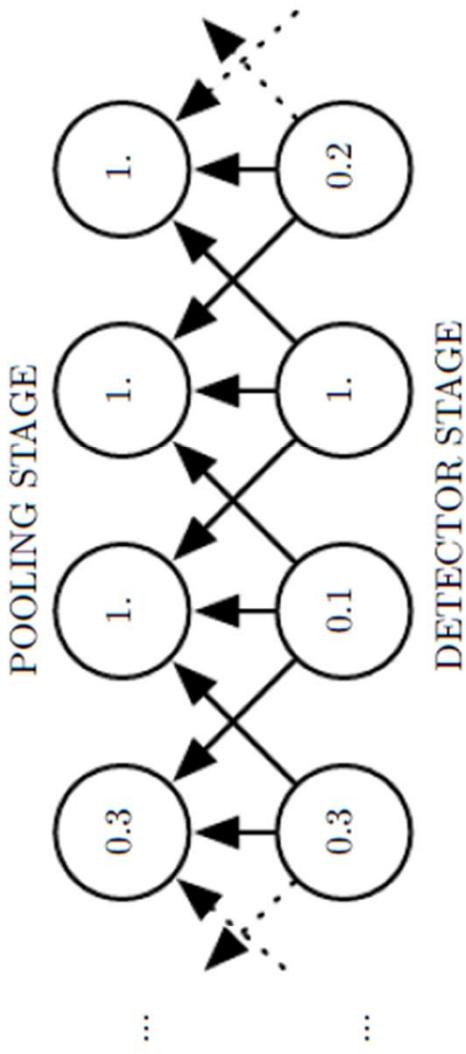
Der Input kommt direkt aus  
einer Schicht, die von Simple-  
Cells gebildet wird.

### Beispiel: **Translationsinvarianz**



### Mögliche Pooling-Operationen

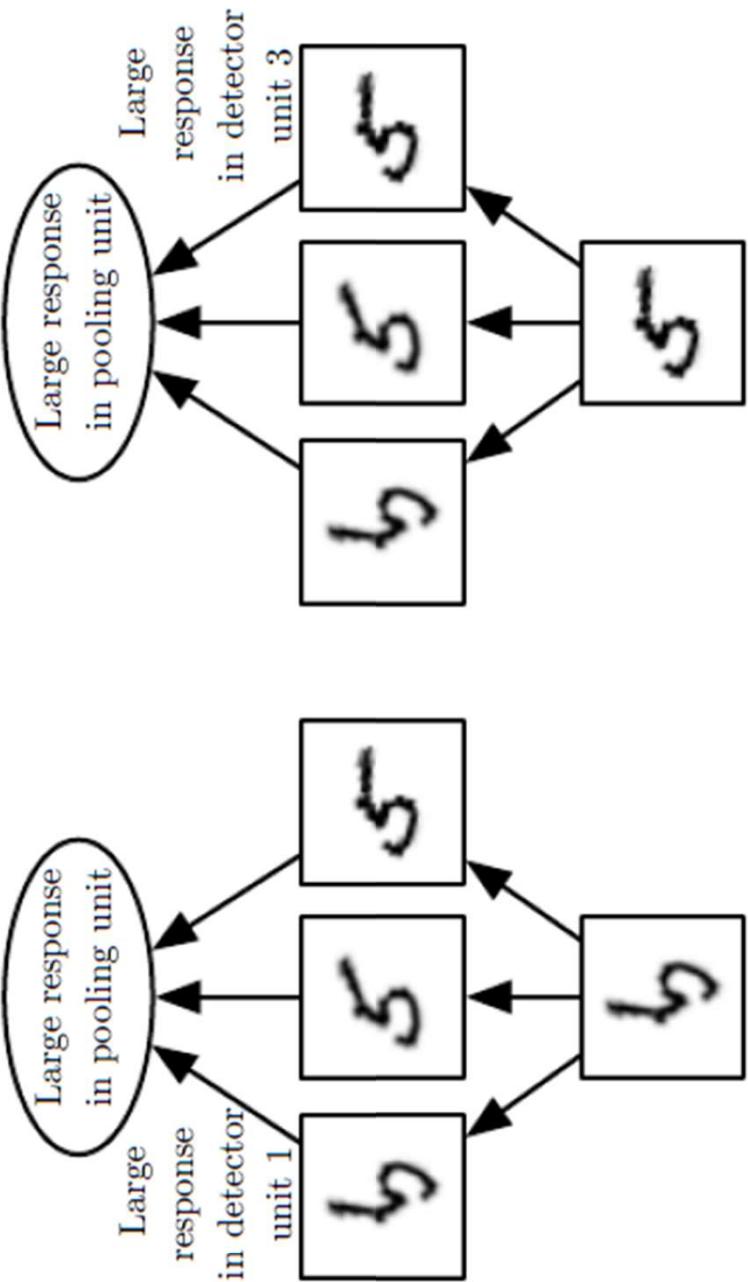
- $\max(\text{RF})$   
Wähle Maximum aus  
dem rezeptiven Feld RF
- $\text{avg}(\text{RF})$   
Mittelung des Inputs aus  
dem rezeptiven Feld RF



**Detector-Stage:** 4 Neuronen ändern ihren Zustand  
**Pooling-Stage:** 2 Neuronen ändern ihren Zustand

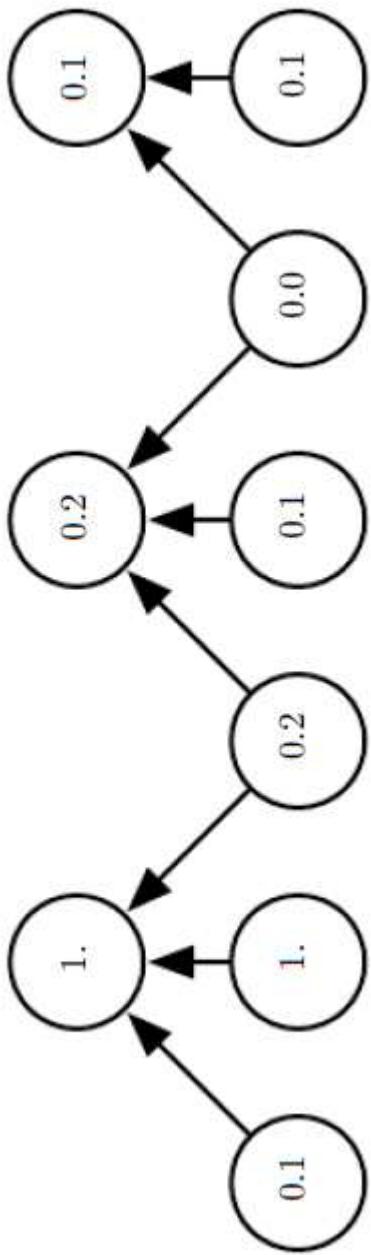
## Beispiel: Lernen von Invarianzen

- Ansichteninvarianz auf höheren Representationsebenen



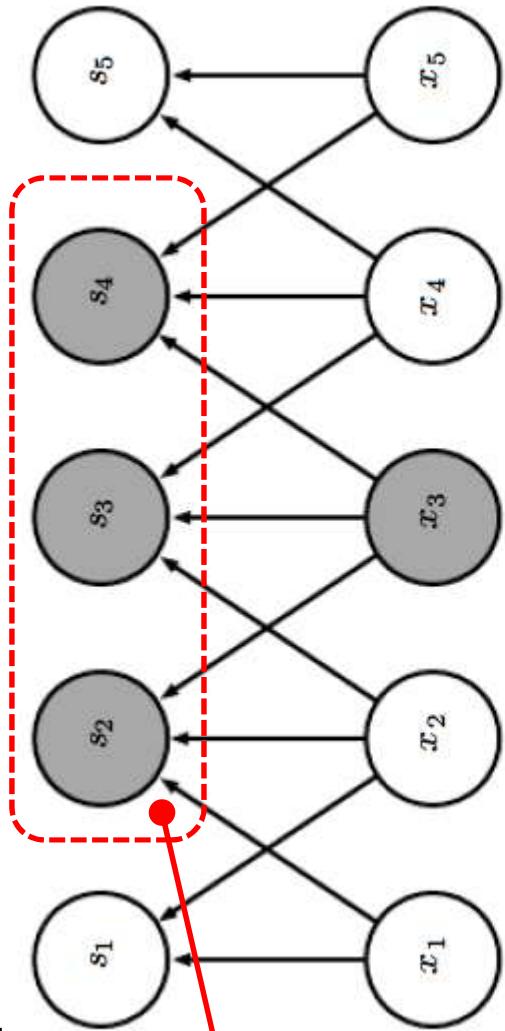
## Beispiel: Pooling mit Unterabtastung („down sampling“)

- Reduktion der Anzahl der Neuronen zur Representation des Inputs



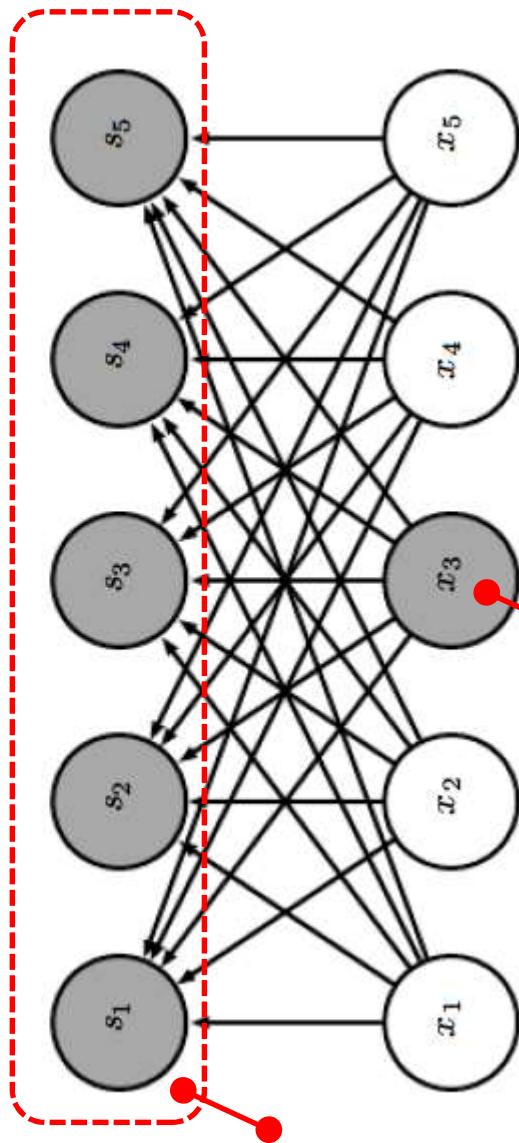
- Spärliche Verbindungen

A.



Einfluß von  $x_3$  bei  
einem Faltungskern

B.



Einfluß von  $x_3$  bei  
vollständige Matrix-  
Multiplikation

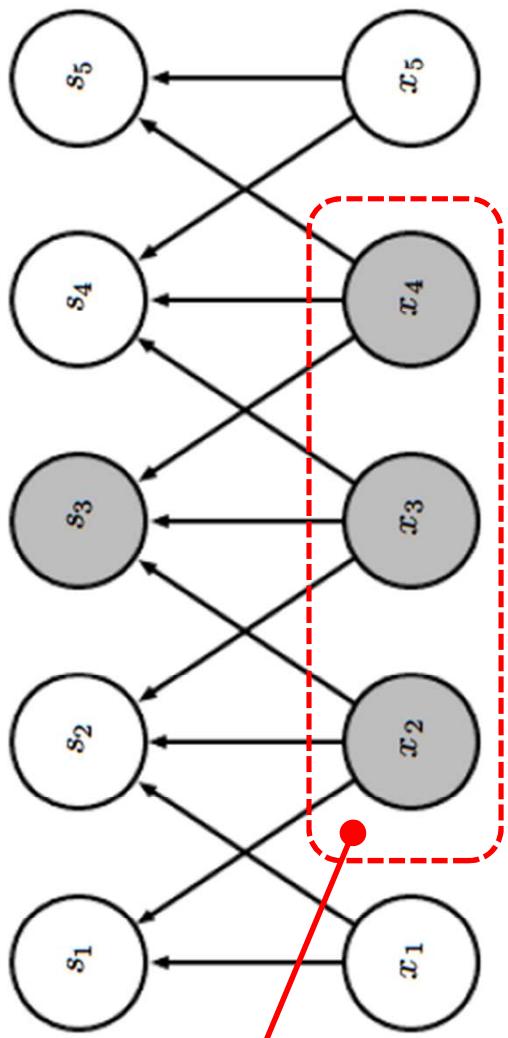
$x_3$ : Input-Neuron

- Spärliche Verbindungen

$s_3$ : Output-Neuron

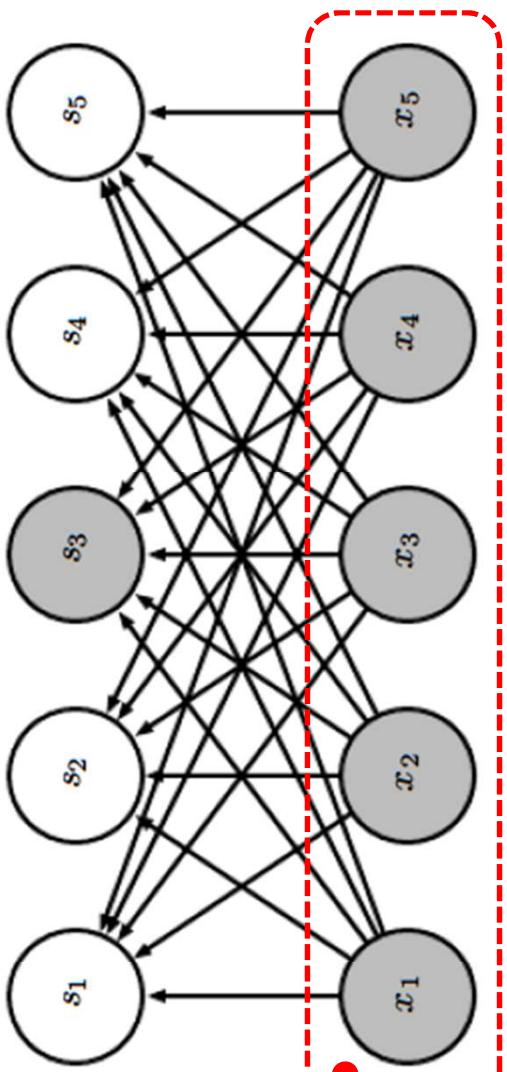
C.

Abhängigkeit von  $s_3$  bei  
einem Faltungskern



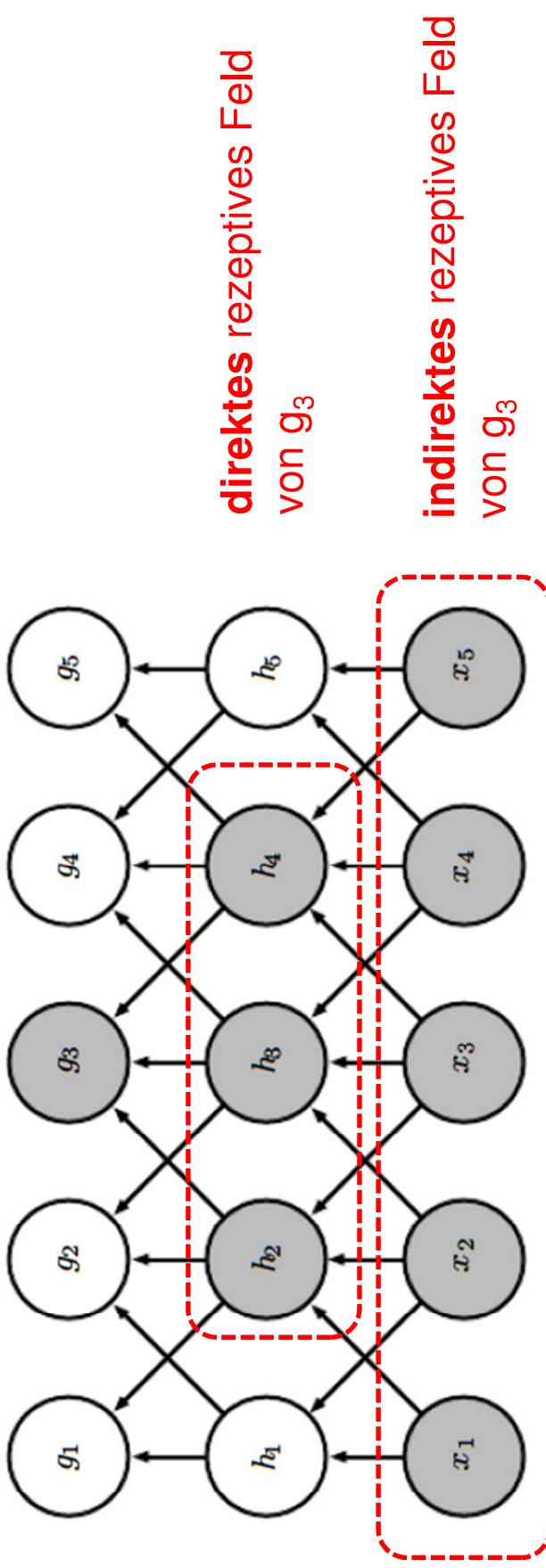
D.

Abhängigkeit von  $s_3$  bei  
vollständige Matrix-  
Multiplikation



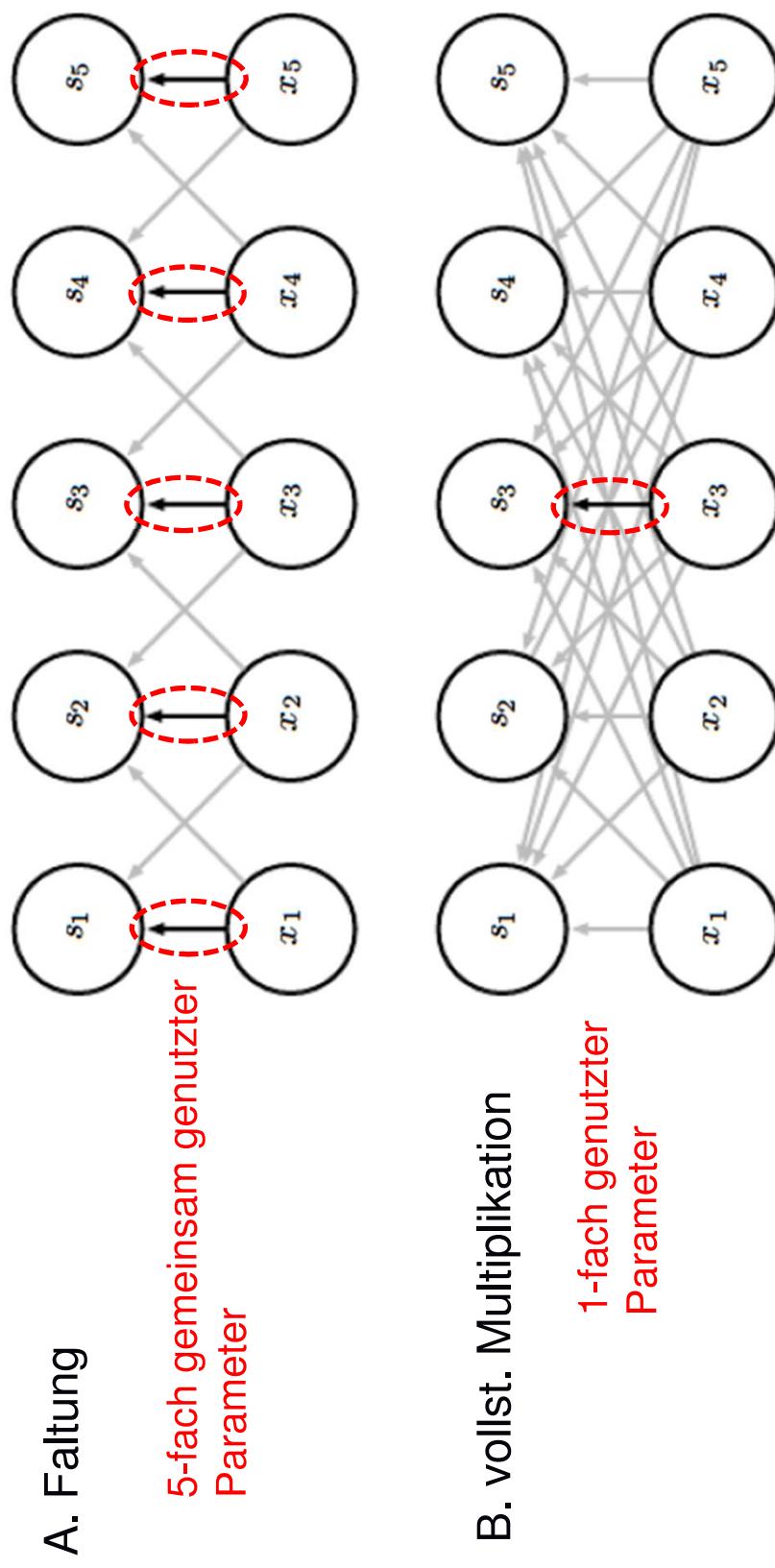
- **Spärliche Verbindungen, vielschichtige Netze („deep Nets“)**

Trotz spärlicher Verbindungen zwischen benachbarten Schichten können Neuronen in tieferen Schichten ein großes rezeptives Feld aufweisen (indirekte Verbindungen)



## • Mehrfachnutzung von Parametern „parameter sharing“

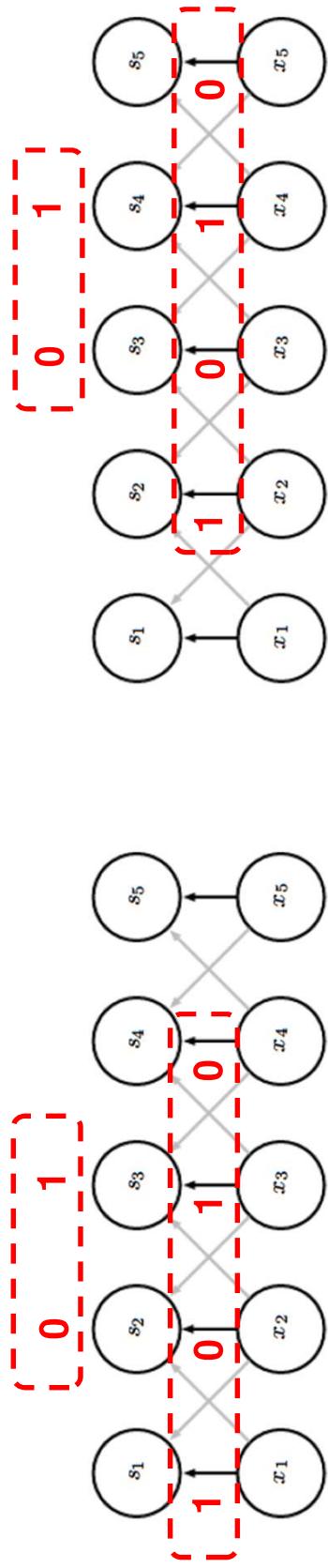
- Ein gegebener Faltungs kern wird in einer Schicht repliziert.  
Dadurch ergibt sich eine Mehrfachnutzung von Parametern



- **Equivarianz bei einer Faltung**

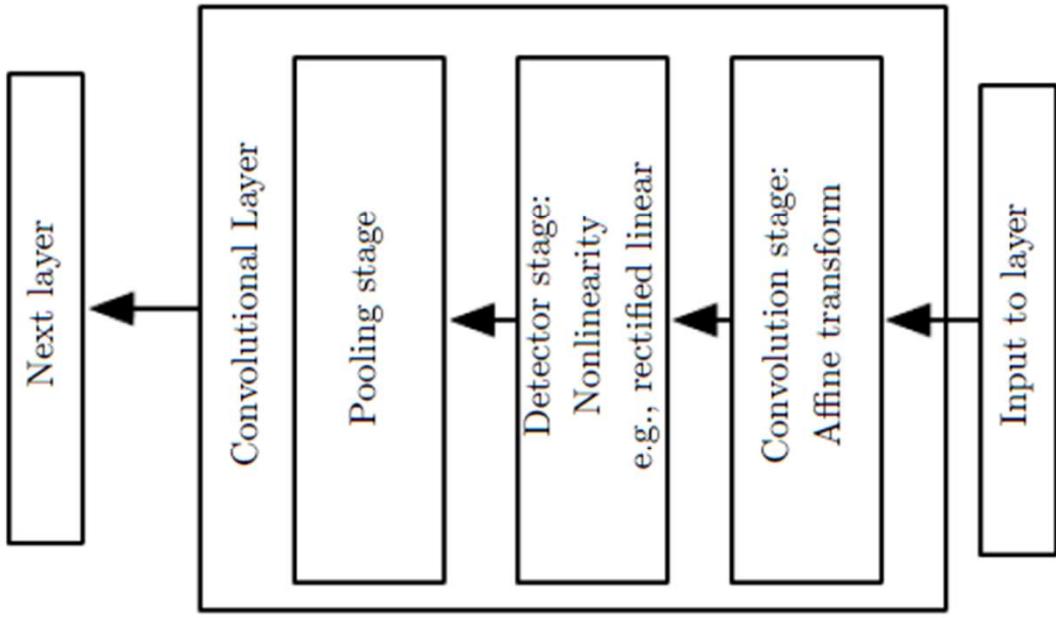
- Eine Verschiebung des Inputs führt zu eben derselben Verschiebung des Outputs

### Beispiel



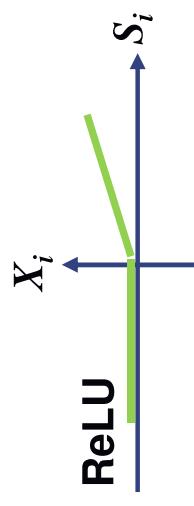
Verschiebung des Inputs um ein Pixel nach rechts bewirkt eine Verschiebung des Outputs um ebenfalls ein Pixel nach rechts.

- Struktur einer Faltungsschicht



$$Y(x, y)$$

$$Y_j = \max(\{X_i\}, i \in RF_j)$$

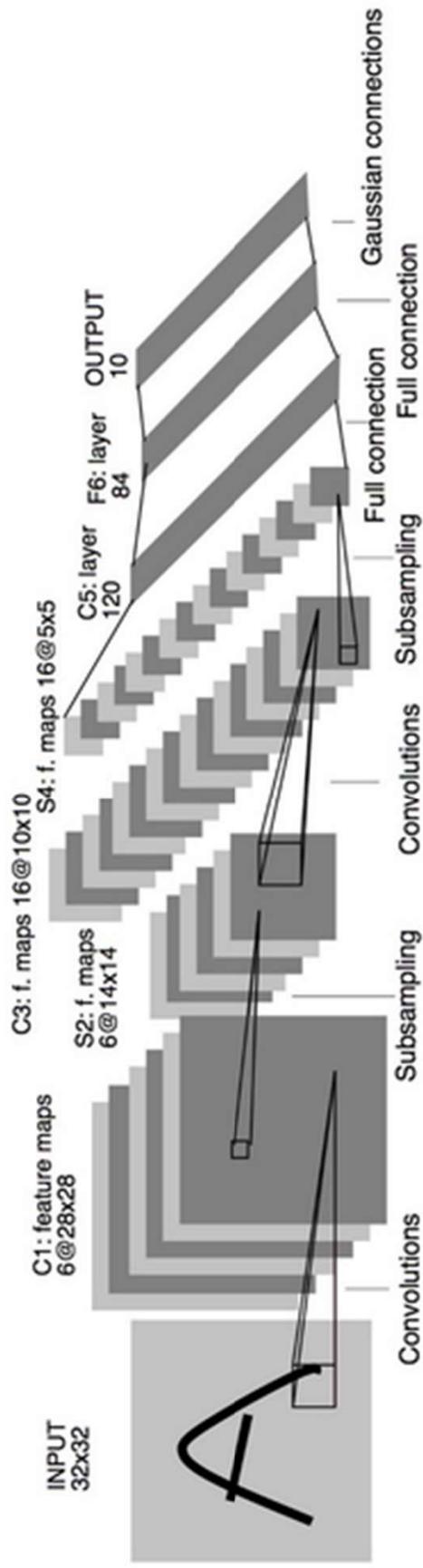


$$S_i(\Delta x, \Delta y) = \sum_{(x,y) \in R(x_0, y_0)} \sum_{W_i(x,y)} I(x-\Delta x, y-\Delta y)$$

$$I(x, y)$$

## 2.9.1 Beispielnetze

### LeNet5 (Ziffernerkennung, MNIST, 60.000 handgeschriebene Ziffern)

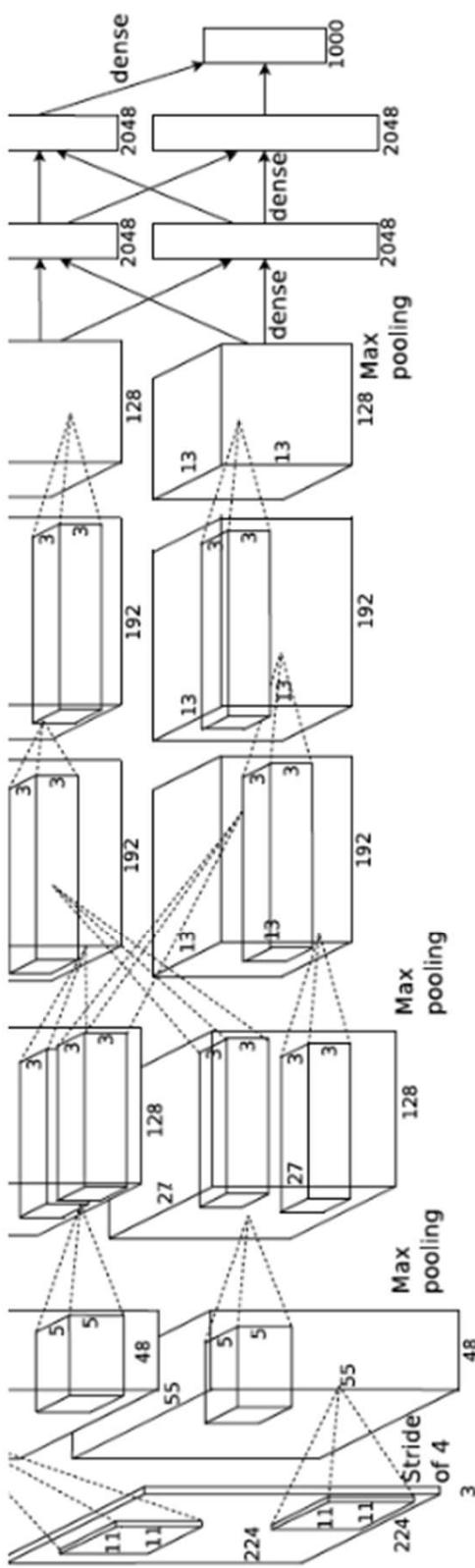


- Faltungsoperation zur Extraktion von räumlichen Merkmalen
- Unterabtastung und Mittelung nicht-überlappender rezeptiver Felder
- $\tanh(\cdot)$  und/oder sigmoid-Funktion zur Modellierung der Aktivierungsfunktion
- MLP als Klassifizierer
- Spärliche Verbindungsmatrizen zur Reduktion der erforderlichen Rechenleistung

LeCun et al.: „Gradient-Based Learning Applied to Document Recognition“, Proc.IEEE 11.1998



## AlexNet, Benchmark: „ImageNet Large-Scale Recognition Challenge ILSVRC“ $14 \times 10^6$ labeled Images, $20 \times 10^3$ Classes



- Ähnlich zu LeNet, höhere Komplexität
- Pooling:  $\max(\text{RF})$  anstelle von  $\text{avg}(\text{RF})$
- ReLU anstelle von  $\tanh(\cdot)$
- 7 verdeckte Schichten,  $650 \times 10^3$  Neurone,  $60 \times 10^6$  Parameter
- GPU Implementation (Beschleunigung gegenüber CPU: Faktor 50)
- Dropout-Regularisierung

A.Krizhevsky et al., „ImageNet Classification with Deep Convolutional Neural Networks“, NIPS 2012

# Clarifai, Benchmark: „ImageNet Large-Scale Recognition Challenge ILSVRC“ $14 \times 10^6$ labeled Images, $20 \times 10^3$ Classes

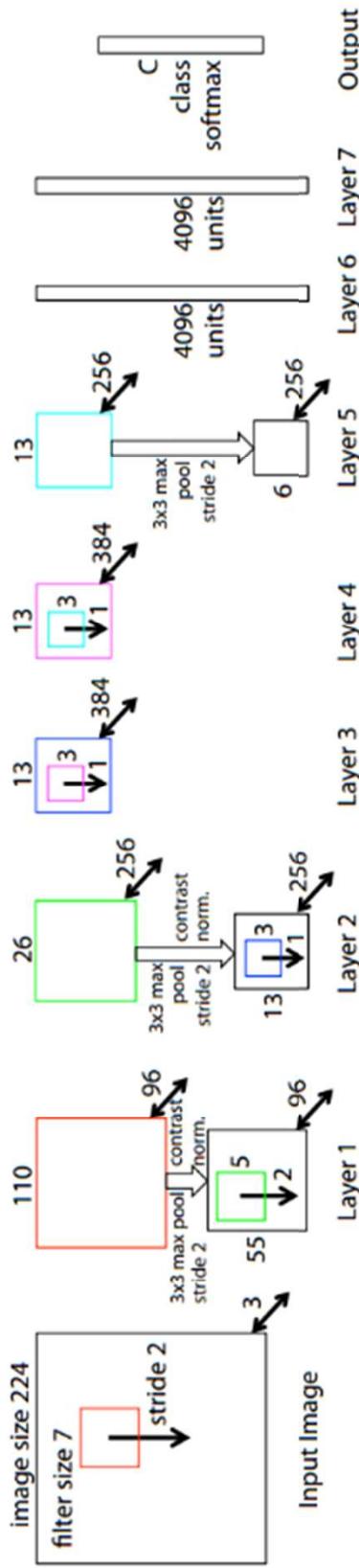


Figure 3. Architecture of our 8 layer convnet model. A 224 by 224 crop of an image (with 3 color planes) is presented as the input. This is convolved with 96 different 1st layer filters (red), each of size 7 by 7, using a stride of 2 in both x and y. The resulting feature maps are then: (i) passed through a rectified linear function (not shown), (ii) pooled (max within 3x3 regions, using stride 2) and (iii) contrast normalized across feature maps to give 96 different 55 by 55 element feature maps. Similar operations are repeated in layers 2,3,4,5. The last two layers are fully connected, taking features from the top convolutional layer as input in vector form ( $6 \cdot 6 \cdot 256 = 9216$  dimensions). The final layer is a  $C$ -way softmax function,  $C$  being the number of classes. All filters and feature maps are square in shape.

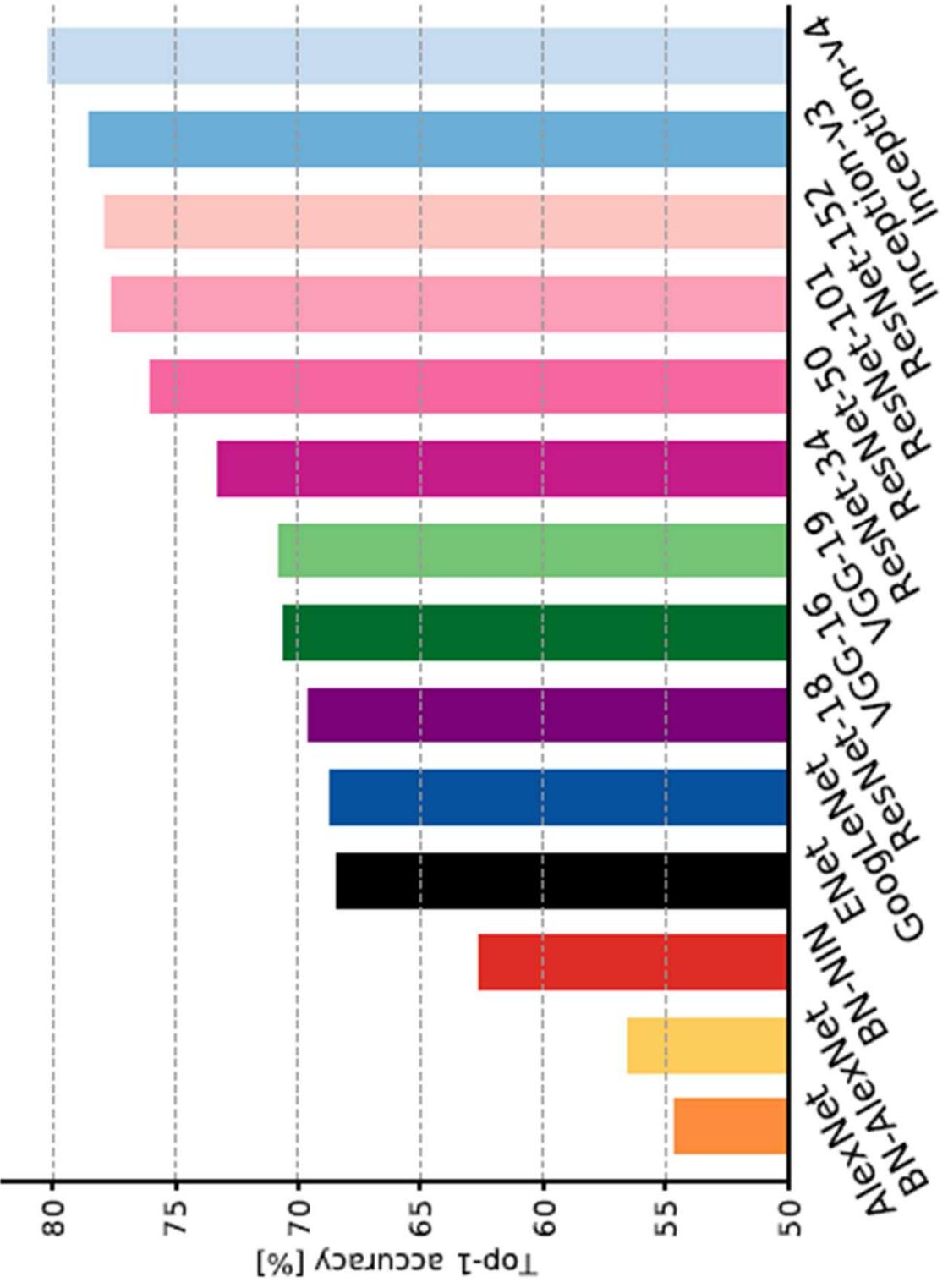
## Zusammenfassung ILSVRC

Team	Year	Place	Error (top-5)	External data
SuperVision – Toronto (AlexNet, 7 layers)	2012	-	16.4%	no
SuperVision	2012	1st	15.3%	ImageNet 22k
Clarifai – NYU (7 layers)	2013	-	11.7%	no
Clarifai	2013	1st	11.2%	ImageNet 22k
VGG – Oxford (16 layers)	2014	2nd	7.32%	no
GoogLeNet (19 layers)	2014	1st	6.67%	no
ResNet (152 layers)	2015	1st	3.57%	
Human expert*			5.1%	

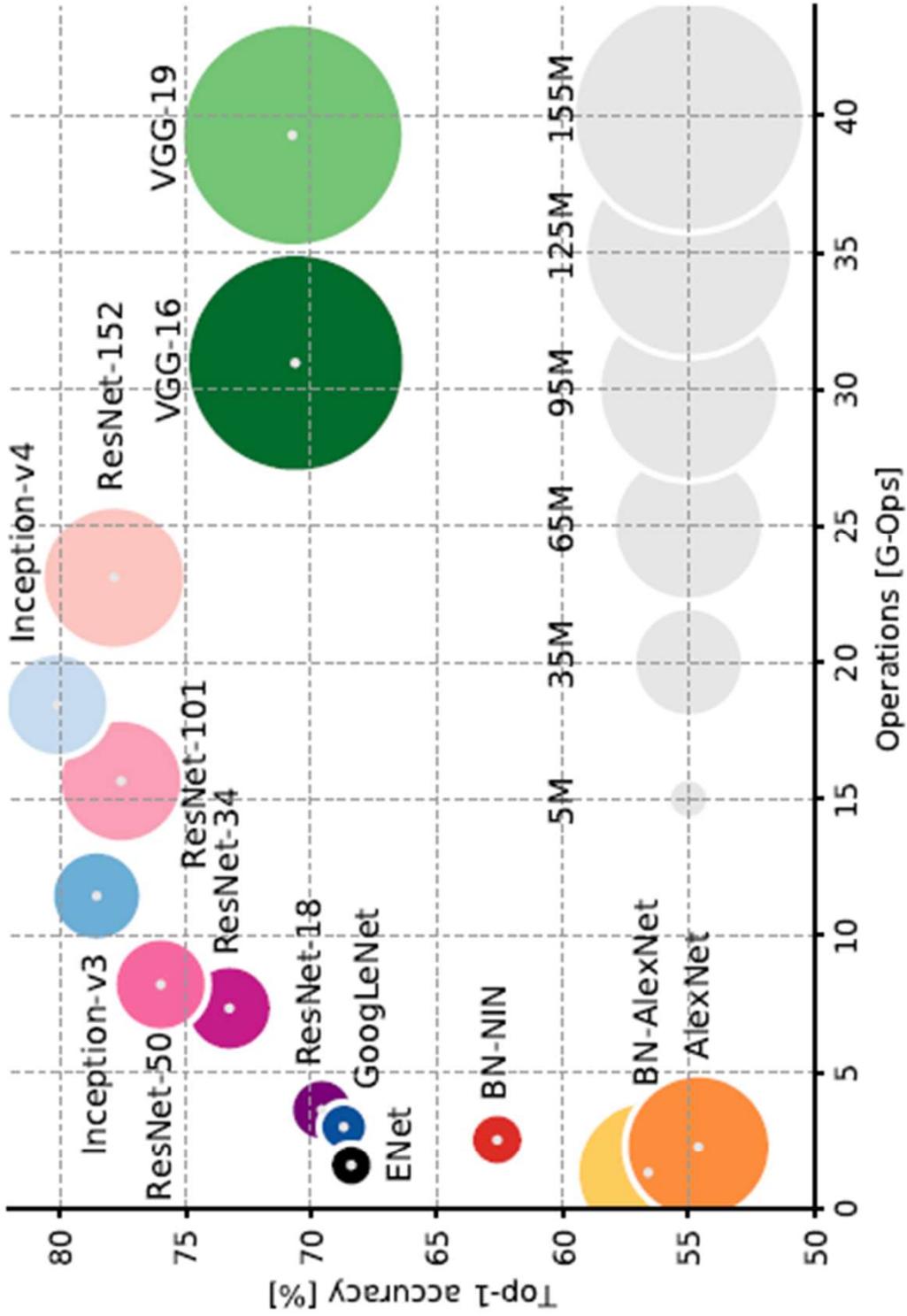
Anm. Top-n accuracy: die korrekte Antwort ist in den n vom Netz höchst-wahrscheinlichen Klassen enthalten.



## Zusammenfassung IL SVRC



## Vergleich: Parameterumfang (Kreisfläche), Accuracy und #Operations



## 2.10 ART Netzwerke, unüberwachtes Lernen

- Adaptive Resonanz Theorie (ART), entwickelt von Stephen Grossberg und Gail Carpenter
- verschiedene Modelle: ART1, ART2, ART3, FuzzyART, ARTMAP
- bei ART Netzen handelt es sich um kompetitive Netze
- ART Netze sind wachsende Netze, d.h. die Zahl der Neuronen kann während des Trainings wachsen. Meist gibt es jedoch eine obere Schranke für die maximale Anzahl zulässiger Neuronen.
- Während des Lernens „erzeugte“ Neuronen können nicht wieder gelöscht werden
- Bei ART1: Eingabedaten und Gewichtsvektoren sind binäre Vektoren



## Das Stabilität-Plastizität Dilemma

### Stabilität

- einmal Gelerntes soll gespeichert bleiben
- das Trainieren irrelevanter oder häufig präsentierter Muster soll vermieden werden, da diese die Klassifikation stark verändern könnten (***blooming, buzzing confusion of irrelevant experience***).

### Plastizität

- Sensitivität für „Neuheit“

Beide Anforderungen konkurrieren: hohe Stabilität *oder* hohe Plastizität

**Es besteht die Notwendigkeit, zwischen Neuheit und Altbekanntem zu unterscheiden, genau so wie zwischen erwarteten und unerwarteten Ereignissen**



## ART 1: Klassifikation binärer Eingabemuster

Die prinzipielle Arbeitsweise des Musterklassifikators ist wie folgt:

- Das Netz bekommt einen Eingabevektor angelegt und versucht ihn, je nach Ähnlichkeit mit gespeicherten Mustern, in eine der vorhandenen Kategorien zu klassifizieren
- Falls der Eingabevektor zu keiner gespeicherten Kategorie ähnlich ist, wird eine neue Kategorie durch Speichern eines dem Eingabemuster ähnlichen Musters erzeugt
- Falls ein Muster gefunden wird, das zum Eingabemuster bei einer vorgegebenen Toleranz ähnlich ist, wird das gespeicherte Muster leicht modifiziert, um es dem neuen Muster ähnlicher zu machen
- gespeicherte Muster, die dem Eingabemuster nicht ähnlich sind, werden nicht verändert



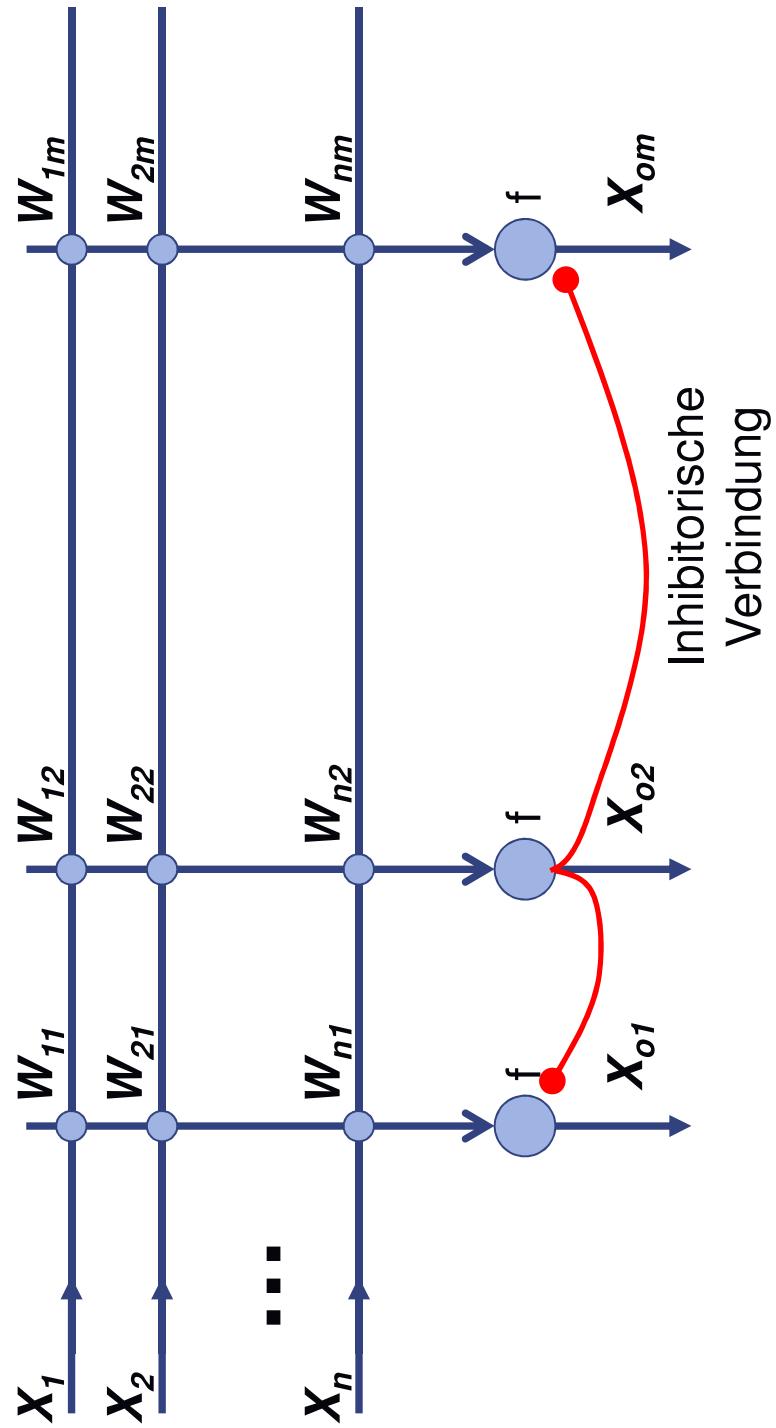
## **ART 1: Klassifikation binärer Eingabemuster**

Damit erhält man eine Lösung des Stabilitäts-Plastizitäts-Dilemmas, denn:

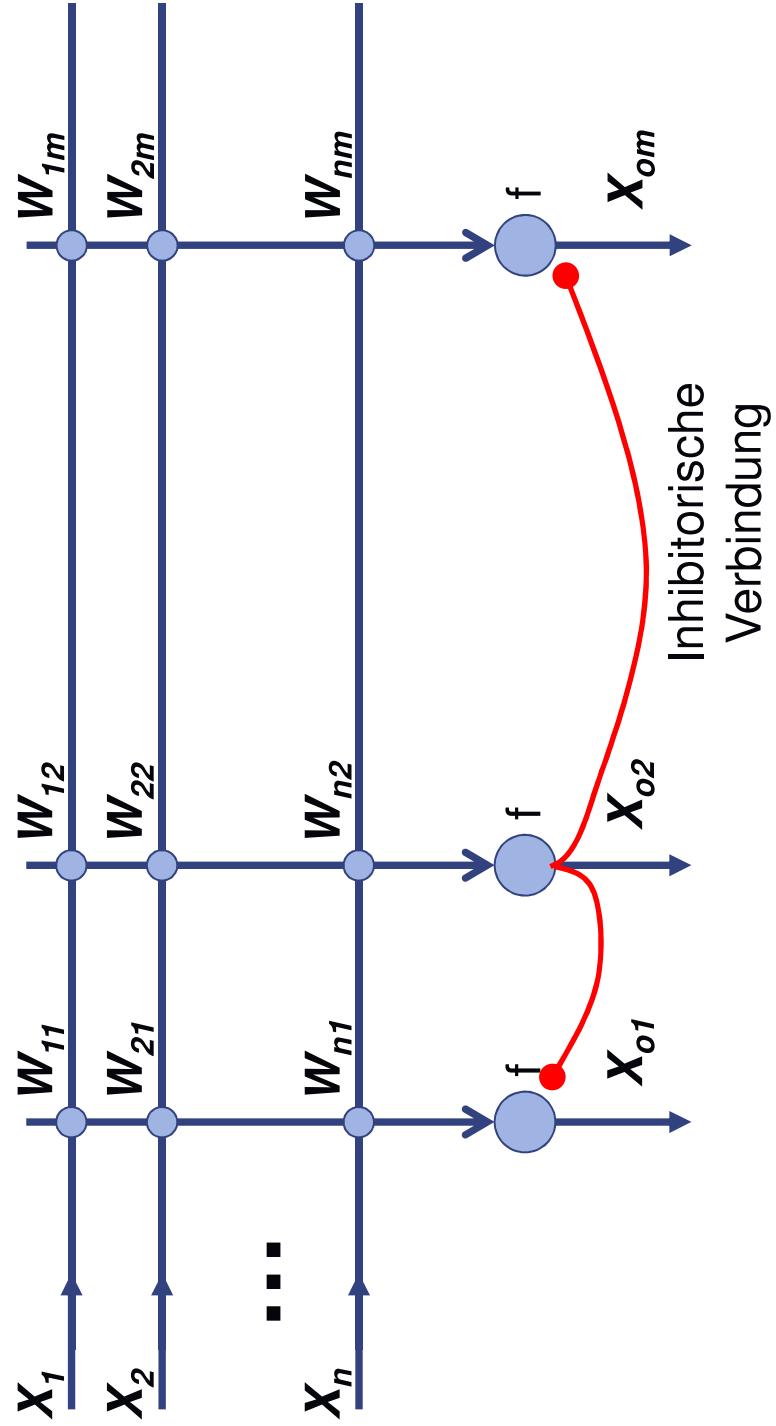
- neue Muster können neue Kategorien erzeugen (Plastizität)
- neue Muster können alte Muster nicht verdecken oder auslöschen, wenn sie diesen nicht sehr ähnlich sind (Stabilität)



## Ansatz: Kompetitives Netz „best match“



Ansatz: Kompetitives Netz „best match“

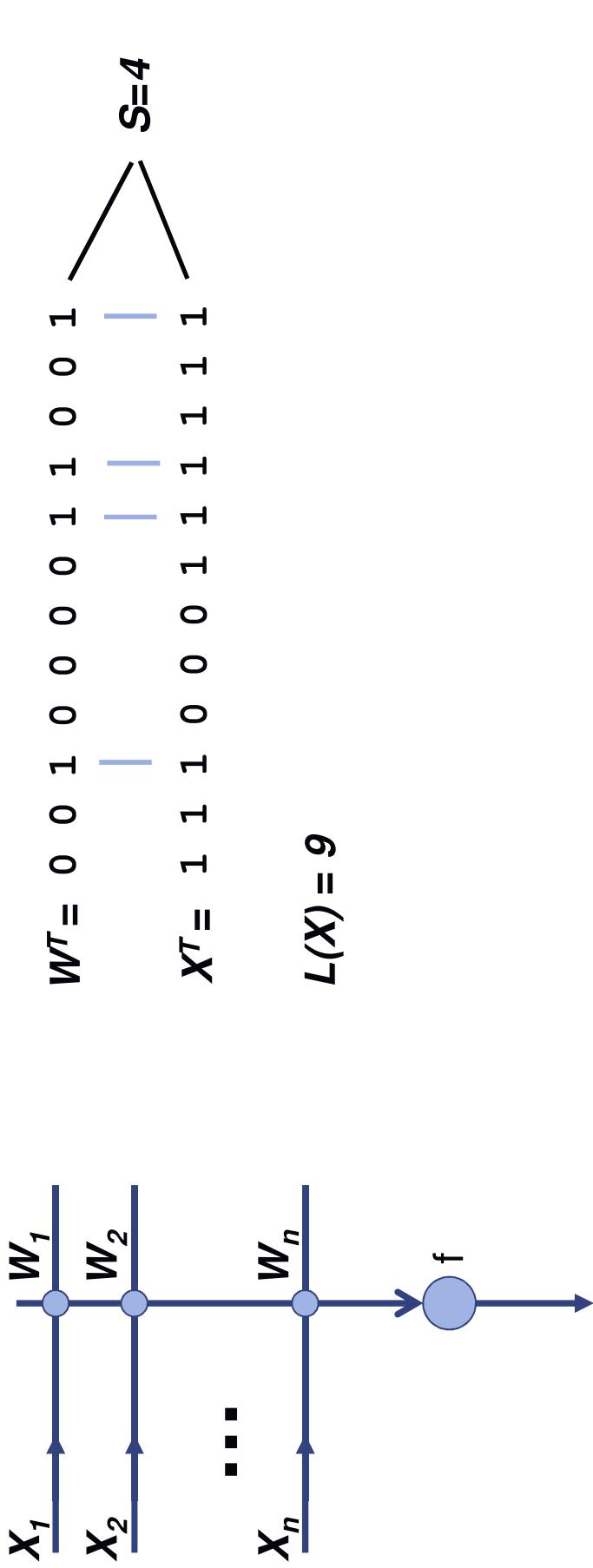


Ansatz: Lernregel (für das jeweilige Gewinnerneuron  $i$ ):

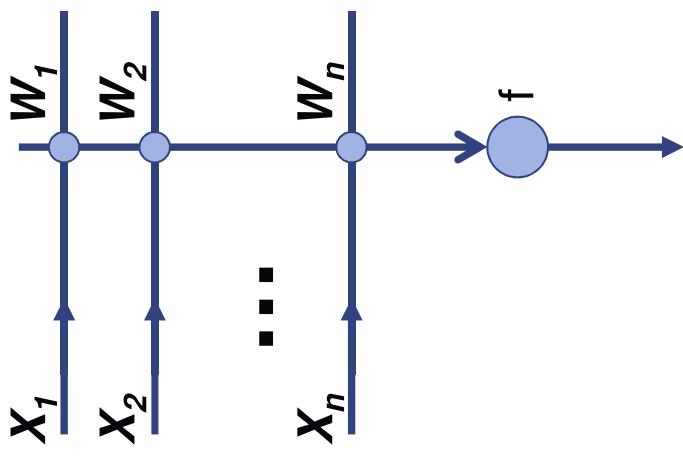
$$W_{ij}(t+1) \leftarrow W_{ij}(t) + \alpha \cdot \left( \frac{X_i}{|X|} - W_{ij} \right)$$

„dreht  $W_j$  in Richtung  $X$ “

## Beispiel: Gewinnerneuron



## Beispiel: Gewinnerneuron



$$W^T = \begin{matrix} 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{matrix}$$
$$X^T = \begin{matrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{matrix}$$

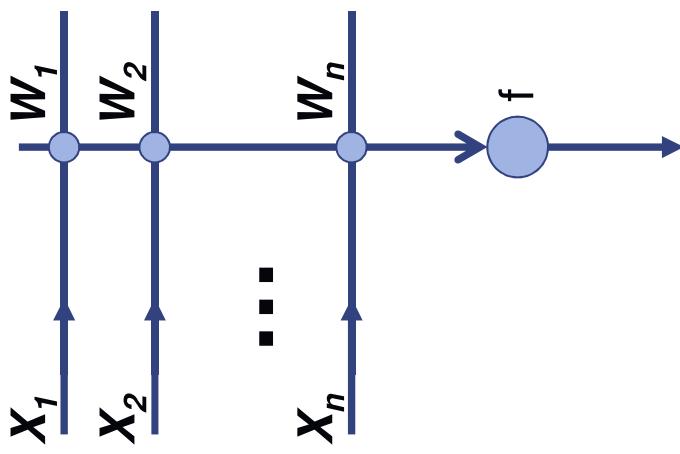
$S=4$

$$L(X) = 9$$

Anwendung der Lernregel (wiederholt)

$$W(t+1) \leftarrow W(t) + \alpha \cdot \left( \frac{X}{|X|} - W(t) \right)$$

## Beispiel: Gewinnerneuron



$$W^T = \begin{matrix} 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{matrix}$$
$$X^T = \begin{matrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{matrix}$$
$$S=4$$

$$L(X) = 9$$

Anwendung der Lernregel (wiederholt)

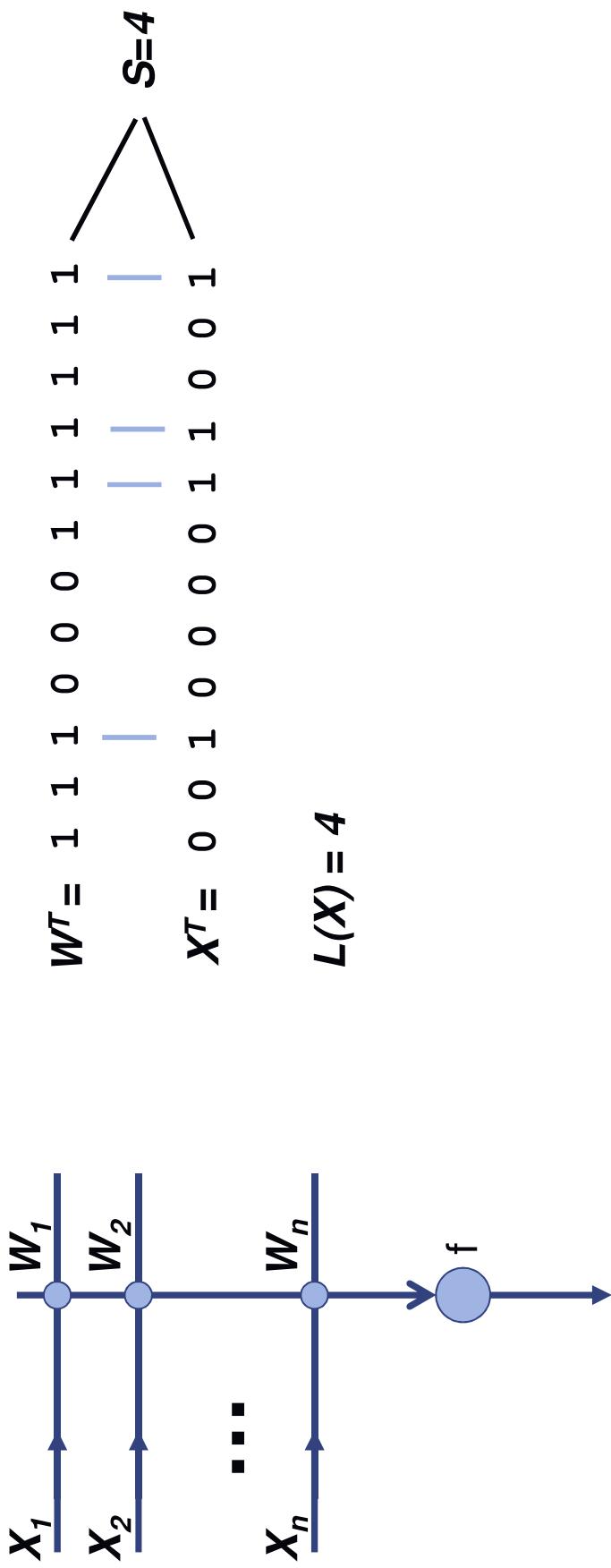
$$W(t+1) \leftarrow W(t) + \alpha \cdot \left( \frac{X}{|X|} - W(t) \right)$$

Führt für große  $t$  auf die Konvergenz

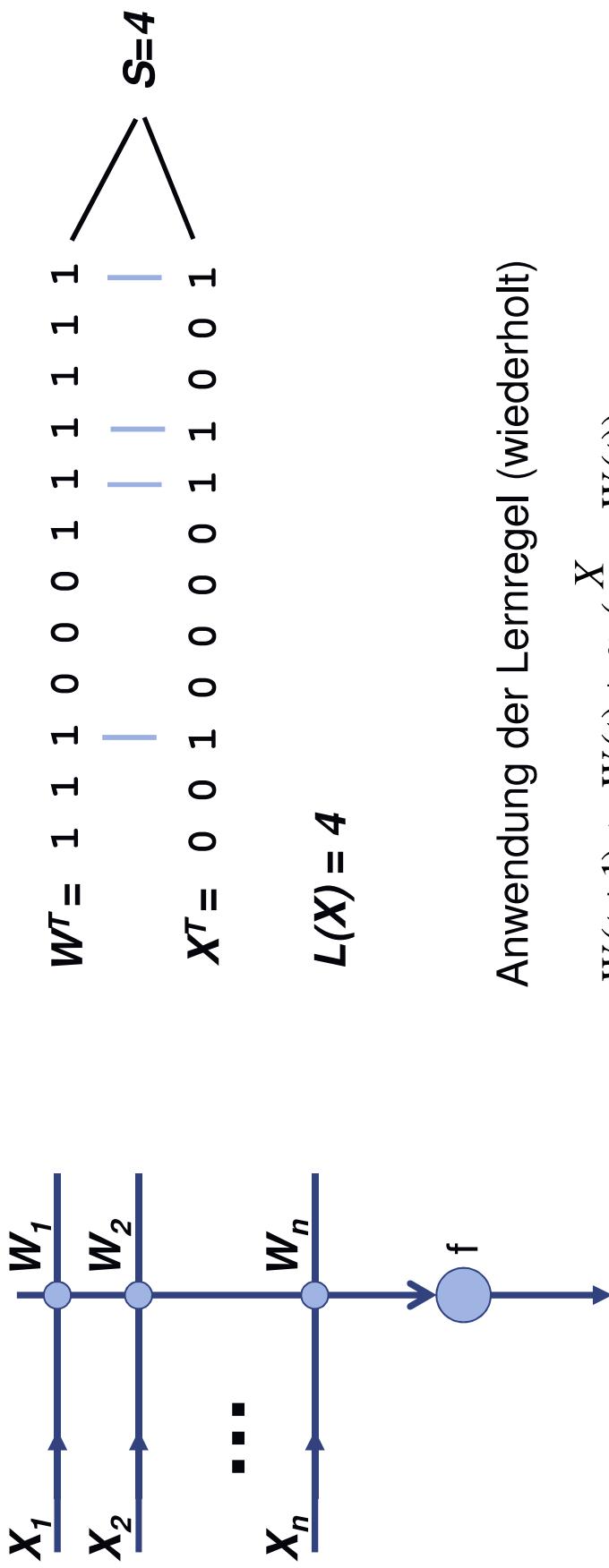
$$W^T \leftarrow \begin{matrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{matrix}$$



## Beispiel: Gewinnerneuron mit geändertem Input



Beispiel: Gewinnerneuron mit geändertem Input



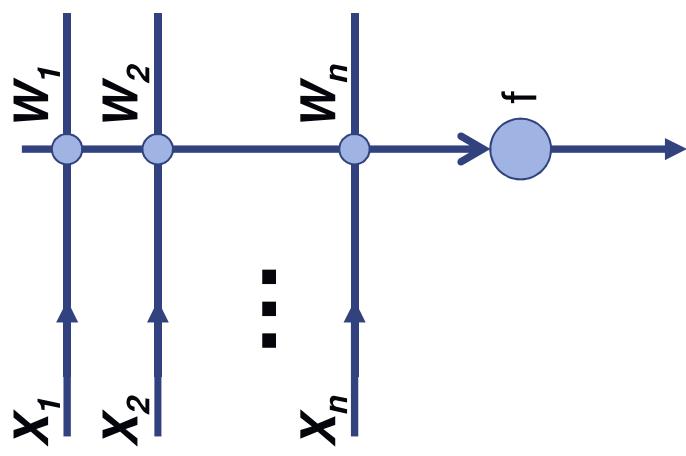
Anwendung der Lernregel (wiederholt)

$$W(t+1) \leftarrow W(t) + \alpha \cdot \left( \frac{X}{|X|} - W(t) \right)$$

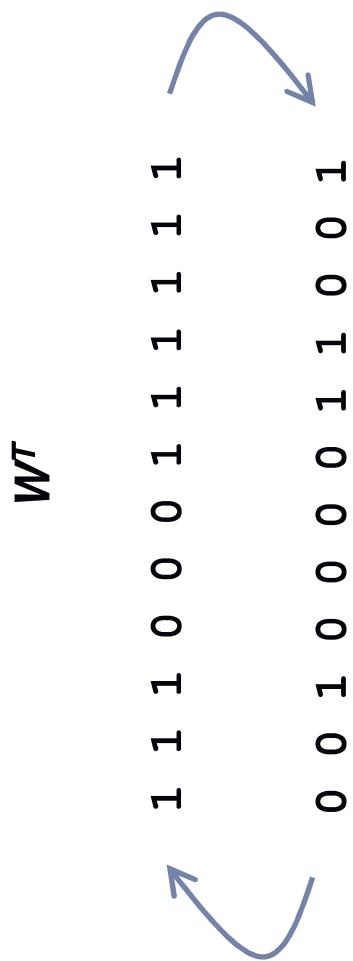
Führt für große  $t$  auf die Konvergenz

$$W^T \leftarrow \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Zusammenfassung:



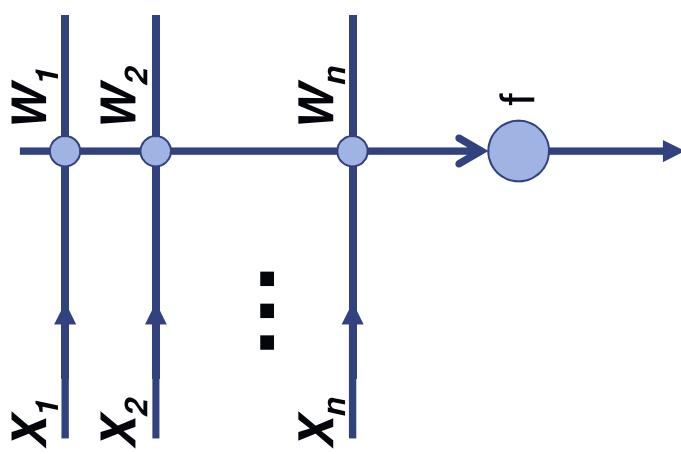
Bei variierendem Input folgt der „repräsentierende Gewichtsvektor“ dem Input (Instabilität)



Grund: Das „Best-Match Neuron“ ist hier dasselbe  
(siehe syn. Potential), da das Skalarprodukt  
kommutativ ist ( $X$  und  $W$  sind vertauschbar)



Abhilfe: Einführung einer Asymmetrie („Ordnung“)



Fall 1:

$$W^\tau = \begin{matrix} 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ & & & | & & | & & & & & \\ & & & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \end{matrix}$$

$X^\tau = \begin{matrix} 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{matrix}$

$L(W) = 9$   
 $S=4$   
 $L(X) = 4$

Fall 2:

$$W^\tau = \begin{matrix} 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ & & & | & & | & & & & & \\ & & & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \end{matrix}$$

$X^\tau = \begin{matrix} 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{matrix}$

$L(W) = 4$   
 $S=4$   
 $L(X) = 9$

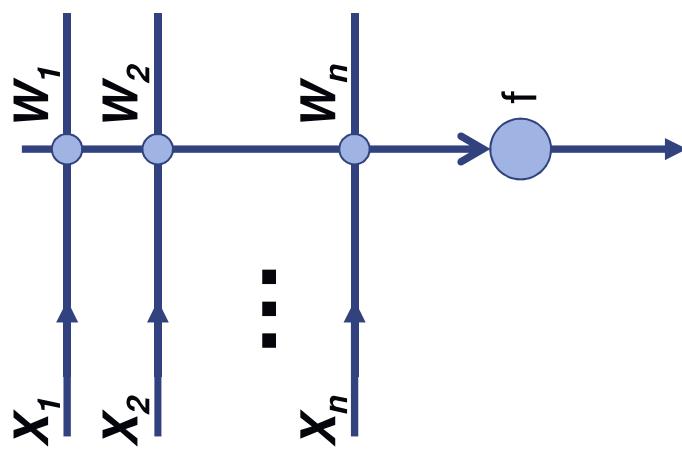
Fall 1 und Fall 2 müssen unterschiedlich behandelt werden, wenn Stabilität gewünscht ist.

Unterscheidung: über  $L(X)$  und  $L(W)$

$L(X)$  ist prinzipiell meßbar.

Problem:

$L(W)$  ist *nicht* meßbar.



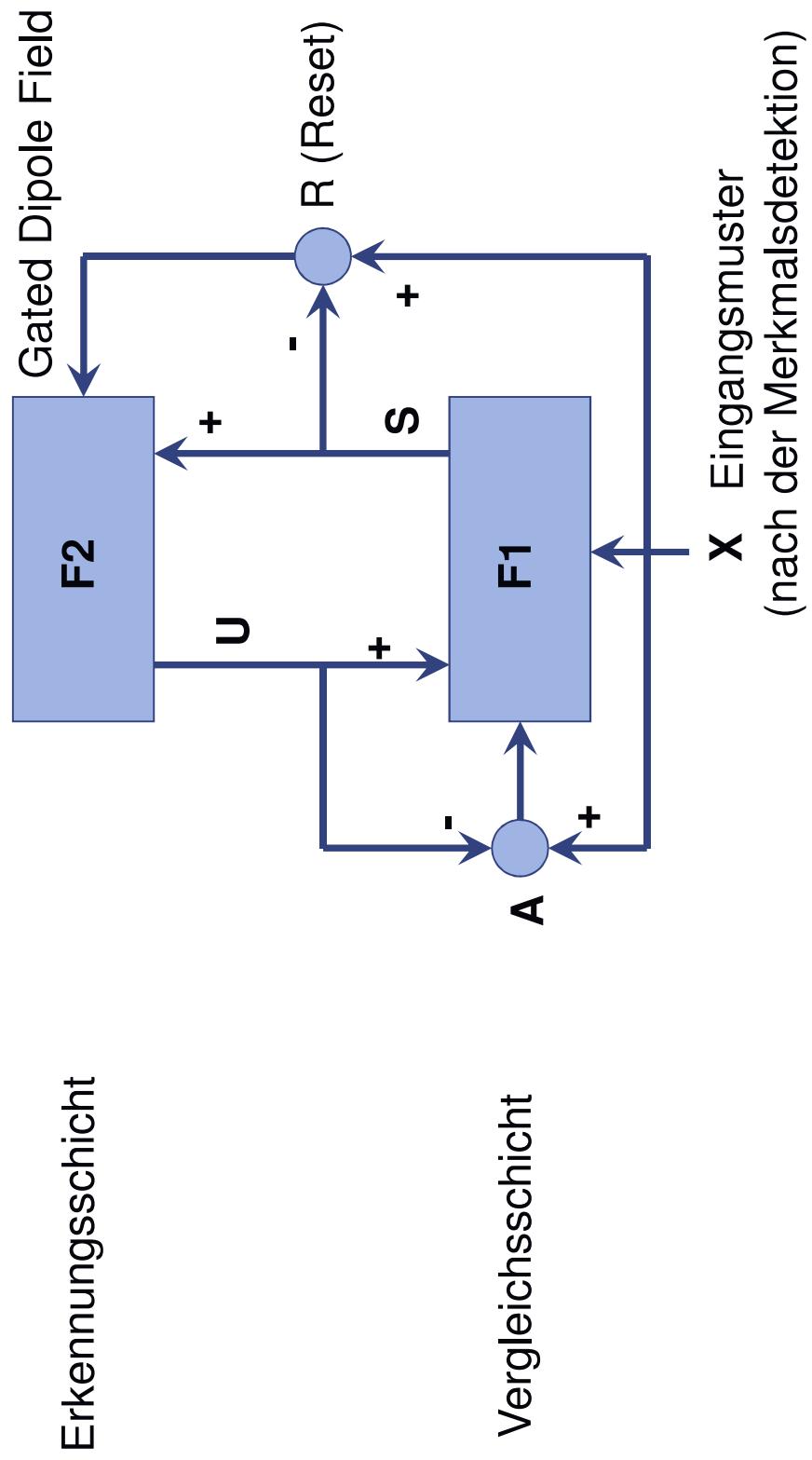
Folge:

Ein **meßbarer** Repräsentant für  $W$  muss generiert werden.

Lösung:

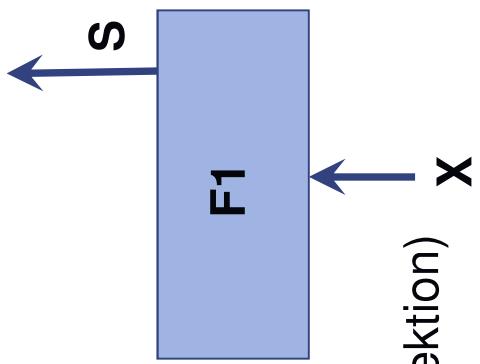
Erzeugung eines „inneren Bildes“ (Leinwand)

## ART 1: Architektur

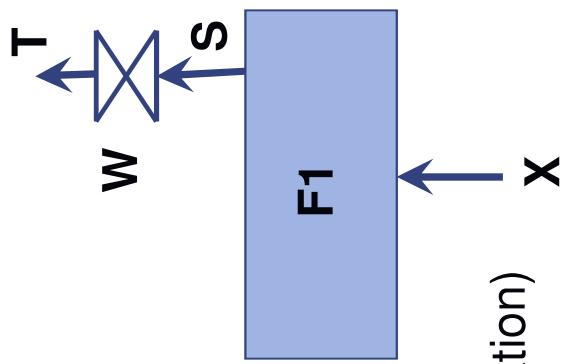


## ART 1: Funktion

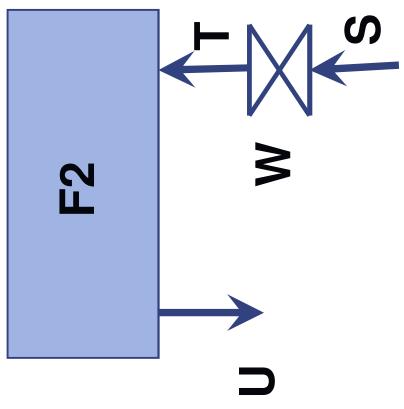
Schritt 1: Eingangsmuster  $X$  erregt Neuronen in  $F_1$ .  
Erzeugung des Musters  $S$  aus  $X$  mit  $S = X$



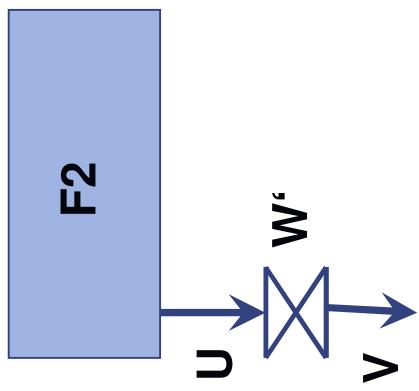
Schritt 2: lineare Abbildung von  $S$  auf das Muster  $T$  durch Multiplikation mit einer Gewichtsmatrix  $W$



## ART 1: Funktion



Schritt 3: Kontrastverstärkung des Musters T durch Schicht F2. Auswahl des Neurons mit der höchsten Aktivität, (Winner-Selektion, Best-Match).  
In F2 wird genau ein Ausgangsneuron aktiv (Muster U)



Schritt 4: lineare Abbildung des Musters U über eine Gewichtsmatrix W auf das Muster V.  
V entspricht einem „erwarteten Muster“, welches nun mit dem Eingangsmuster X verglichen wird.

Stimmen V und X hinreichend überein, hat das System das Eingangsmuster erkannt („Erwartung“ und Input passen zusammen). Gleichzeitig sind die gespeicherten Repräsentationen für X in W und W' hinreichend ähnlich, so dass eine nachfolgende Gewichtsadaption W und W' nur minimal stört.



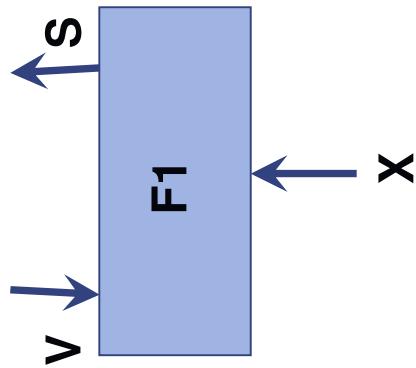
## ART 1: Funktion

Stimmen  $V$  und  $X$  nicht überein, so muss das System versuchen, eine anderes „erwartetes“ Muster  $V$  aus den bereits gespeicherten Mustern in  $W$  und  $W'$  zu erzeugen. Hierfür sind zwei Schritte notwendig.

- (I) Quantifizierung, wie ein Vergleich durchzuführen ist
- (II) Langanhaltende Inhibition des aktiven U-Neurons, um andere Repräsentationen in  $W'$  abrufen zu können.

Schritt 5: Generierung eines Musters  $S$  aus dem Vergleich von  $X$  und  $V$ . Gebildet wird  $S$  aus der komponentenweisen UND-Verknüpfung zwischen  $X$  und  $V$ :

$$S \leftarrow X \cdot V$$



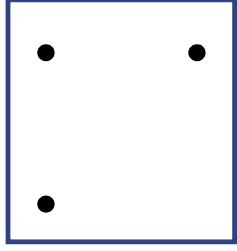
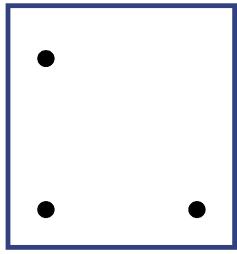
$S$  enthält jetzt den gemeinsamen Überlapp von  $X$  und  $V$ . Aus dem Vektor  $S$  und  $X$  wird nun bestimmt, ob  $V$  und  $X$  ähnlich sind. Die Ähnlichkeit bestimmt sich aus dem Betragsverhältnis von  $X$  und  $S$ .



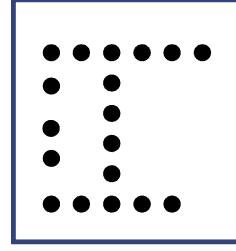
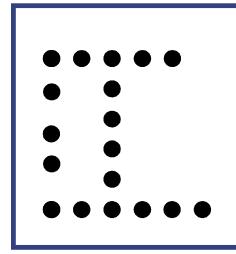
## ART 1: Funktion

Begründung für die Wahl des Betragsverhältnisses von X und S als Ähnlichkeitsmaß:

Betrachte hierzu zwei Vektoren, mit einem gemeinsamen Überlappbereich und nicht-überlappenden Komponenten (hier: Merkmale).



Muster 2



In beiden Fällen unterscheiden sich Muster 1 und Muster 2 um die selben beiden Merkmale. Während in Fall 2 die Muster 1 und Muster 2 ähnlich erscheinen (durch die Vielzahl übereinstimmender Merkmale), erscheinen in Fall 1 die Muster eher unterschiedlich (durch die relativ geringe Anzahl gemeinsamer Merkmale, im Verhältnis zu den sich unterscheidenden Merkmalen)

## ART 1: Funktion

Es ist sinnvoll, als Ähnlichkeit sowohl den gemeinsamen Überlapp als auch die nicht-überlappenden Merkmale heranzuziehen. Zwei Muster S und X gelten als ähnlich, wenn

$$\frac{|S|}{|X|} \geq \sigma \quad 0 < \sigma < 1$$

erfüllt ist. Mit Hilfe dieser Definition ist es z.B. möglich, neue Muster zu detektieren, was z.B. unter Verwendung eines Skalarproduktes nicht einfach möglich ist. Wir zeigen dies in den nachfolgenden Überlegungen.

**Skalarprodukt:** Bei der Abbildung von S nach U (Schritt 1 – Schritt 3) wird ein anderes Ähnlichkeitsmaß verwendet: das Skalarprodukt. Z.B. bei rein binären Vektoren ergibt das Skalarprodukt gerade den gemeinsamen Überlapp zweier Muster. Über die nicht-überlappenden Merkmale oder gar das Verhältnis nicht-überlappende zu überlappende Merkmals liefert das Skalarprodukt von Binärvektoren zunächst keine Aussage.

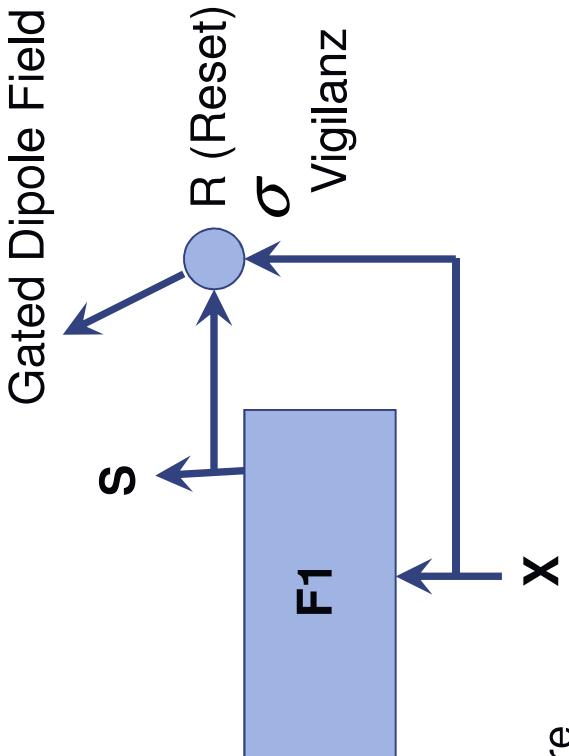
Wir stellen uns nun zunächst die Aufgabe, mit Hilfe eines Skalarproduktes zwei Muster von einander zu trennen, wobei das eine Muster in dem zweiten Muster vollständig enthalten ist. Die Trennung soll durch zwei Gewichtsvektoren W1 und W2 erfolgen, die jeweils als Kopie der zu trennenden Muster mit spezieller Normierung aus den Mustern P1 und P2 abgeleitet sind. Die Trennung erfolgt mit einem competitiven Netzwerkes (Suche des max. Skalarprod.).



## ART 1: Funktion

Um dem System das Lernen des zweiten Gewichtsvektors zu ermöglichen, muß ein weiteres Ähnlichkeitsmaß herangezogen werden, welches an anderer Stelle im Signalfluß ausgewertet wird. Wie bereits angedeutet, ist es im ART System sinnvoll, aus dem erwarteten Vektor  $V$  die Ähnlichkeit zur Eingabe  $X$  abzuleiten und ggf. „Neugigkeit“ (engl. Novelty) zu detektieren. Ähnlichkeit ist an dieser Stelle gegeben, falls

$$\frac{|S|}{|X|} \geq \sigma \quad 0 < \sigma < 1$$



erfüllt ist.

Schritt 6: Bestimmung der Ähnlichkeit mit Hilfe des Neurons  $R$ , welches das neue Ähnlichkeitsmaß auswertet. Sind  $X$  und  $S$  unähnlich, feuert das Neuron  $R$  und inhibiert das gerade aktive (Winner-) Neuron in  $F2$  über die Dipol-Felder für lange Zeit. Dadurch kann ein anderes Neuron ausgewählt werden und andere Signale (Muster) können die Schleife neu durchlaufen.

## ART 1: Funktion

Sind alle Neuronen der F2-Schicht über die **Gated Dipole Fields** deaktiviert, liegt offensichtlich ein neues Muster vor. Das ART System legt dann ein neues Teilsystem an, welches sich in einem weiteren Neuron in F2 und je einem neuen Gewichtsvektor in  $W$  und  $W'$  widerspiegelt.

Generierung von S aus X und V:

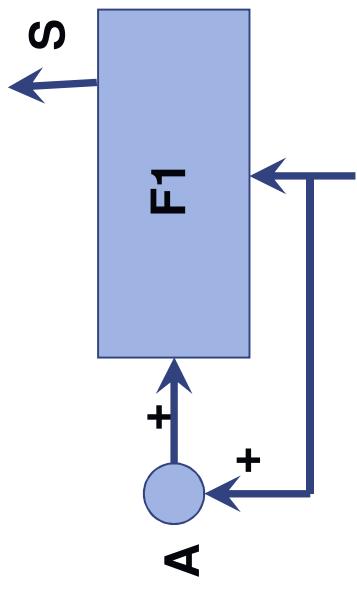
Wird der Eingangsvektor  $X$  dem System präsentiert, darf es nicht passieren, dass das System diesen Vektor als „neu“ erkennt. Das System kann zu diesem Zeitpunkt noch keinen Vektor  $V$  erzeugen, damit muss die „Verundung“ in  $F1$  deaktiviert werden. Um Diese Phase von anderen Phasen zu unterscheiden, wird ein weiteres Neuron  $A$  eingefügt, welches auf die  $F1$ -Schicht in besonderer Weise wirkt. Der Ausgang von Neuron  $A$  ist mit jedem Neuron der Schicht  $F1$  verbunden. Folgende Fälle sind zu unterscheiden:



## ART 1: Funktion

(I) Muster X wird angelegt, V ist nicht präsent

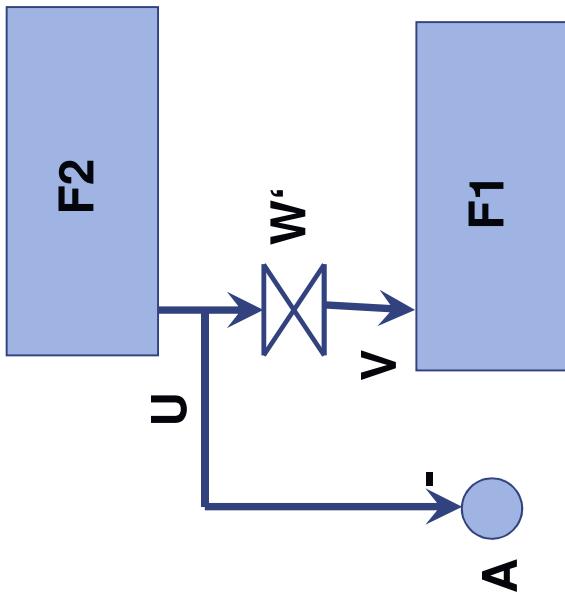
Neuron A generiert für diesen Fall ein positives Signal für F1, um das Signal X zu verstärken



Eingangsmuster X  
(nach der Merkmalsdetektion)

(II) Spontane Aktivität von F2

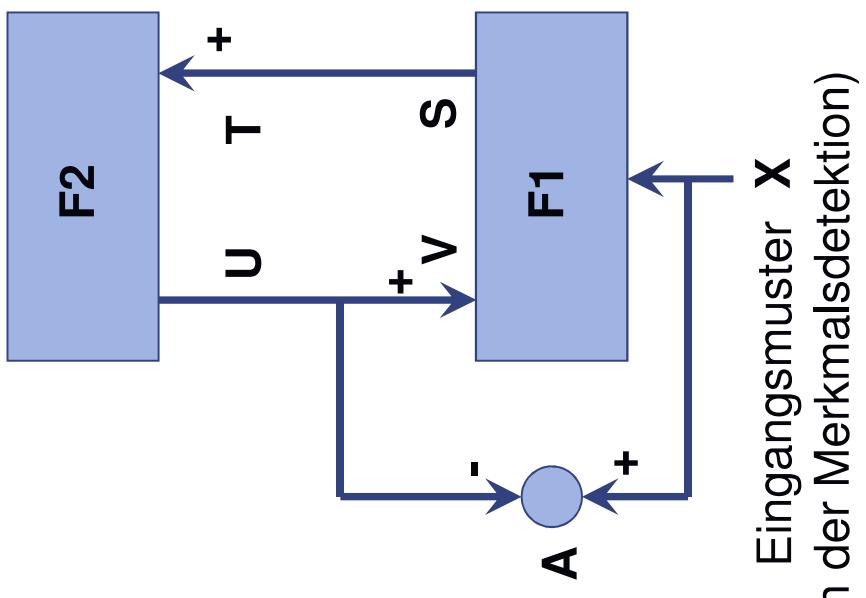
Neuron A generiert für diesen Fall kein Signal für F1, wodurch weder X noch V\*X verstärkt werden



## ART 1: Funktion

(III) Muster X wird angelegt, V ist präsent

Neuron A generiert für diesen Fall kein Signal für F1



Eingangsmuster **X**  
(nach der Merkmalsdetektion)

Zusammenfassung (I)-(III)

Im Fall (I) erhält jedes Neuron von A und von jedem aktiven Part aus X positiven Input

Im Fall (II) erhält jedes Neuron nur von einem aktiven Part aus V einen positiven Input

Im Fall (III) erhält jedes Neuron positiven Input von den aktiven Komponenten aus X und V



## ART 1: Funktion

Die 2/3 Regel:

Ein Neuron aus der Schicht F1 feuert, wenn mindestens zwei (von Dreiern) Eingänge aktiv sind. Somit kann mit Hilfe von Neuron A das initiale Anlegen von X und die Such-Phasen unterschieden werden.

$$S_i = \begin{cases} 1 & X_i \cdot V_i \vee X_i \cdot A \vee V_i \cdot A = 1 \\ 0 & sonst \end{cases}$$



## Adoptionsphase:

Adaptiert werden Gewichte, wenn entweder ein neues Muster gefunden wurde und neue Gewichtsvektoren angelegt werden, oder aber wenn ein bekanntes Muster gefunden wurde. Adaptiert werden jeweils nur Gewichte, die zum gefundenen Muster oder zu den neu angelegten Gewichtsvektoren gehören.

$$\tau \cdot \frac{dW_{ij}}{dt} = \left[ (1 - W_{ij}) \cdot D - W_{ij} \cdot \left( \sum_{k=1}^n X_k - 1 \right) \right] \quad (\text{for } W \text{ and } W')$$

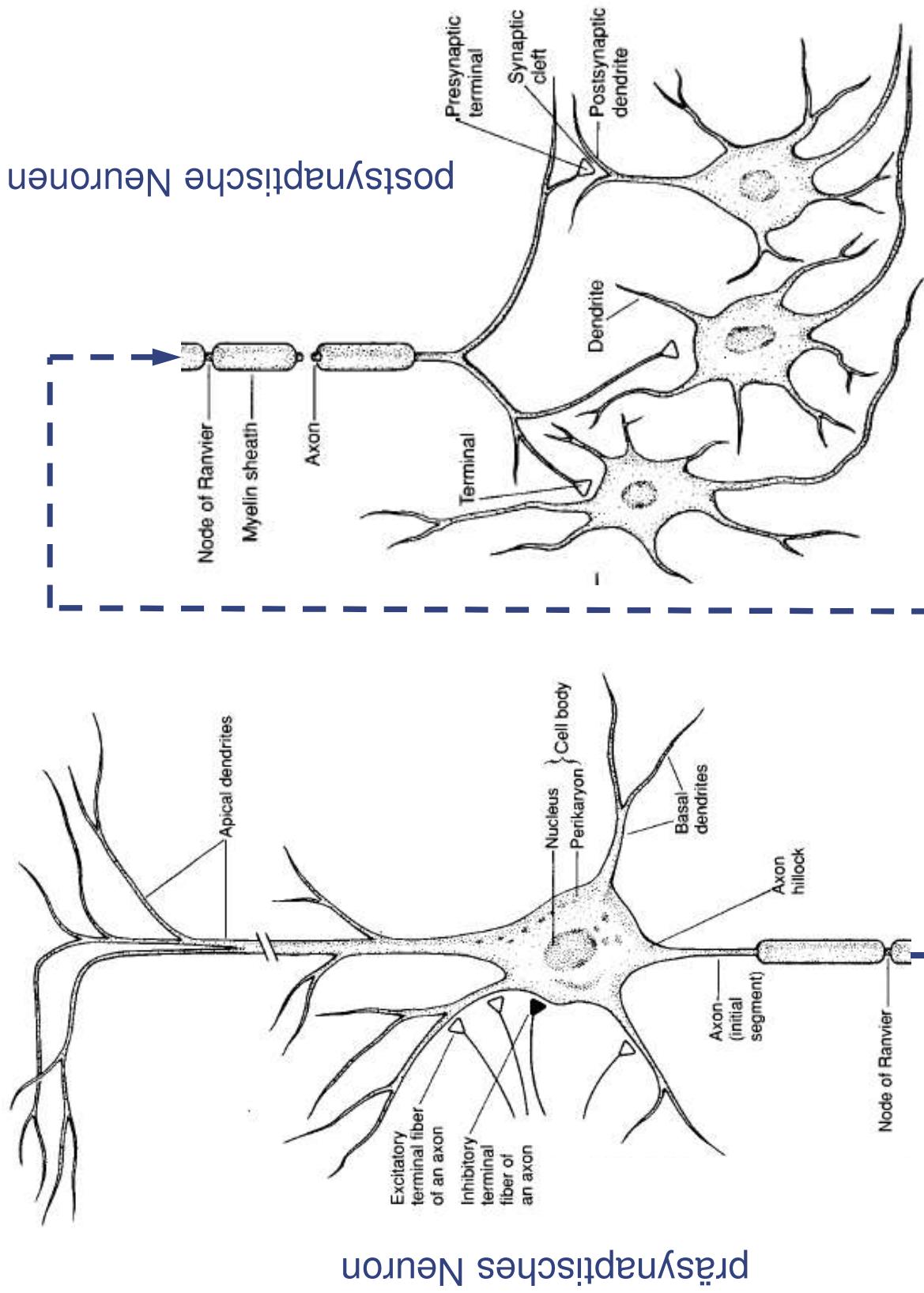
**(for the winning neuron  $U_j$  and  $X_i=1$ )**

$$\tau \cdot \frac{dW_{ij}}{dt} = -W_{ij} \cdot \left( \sum_{k=1}^n X_k \right) \quad D = 2 \dots 3 \quad (\text{for the winning neuron } U_j \text{ and } X_i=0)$$

### **3: Biologische Neuronale Netze**



### 3.1 Neurone



**Neuron:** Nervenzelle (griech. νεῦρον, *neūron* :Nerv)

Begriff geprägt von Heinrich Wilhelm Waldeyer 1881

**Synapse:** Kontaktstelle zwischen Nervenzellen  
(griech. *syn* = zusammen, ἄπτειν, *hapttein* = klammern )

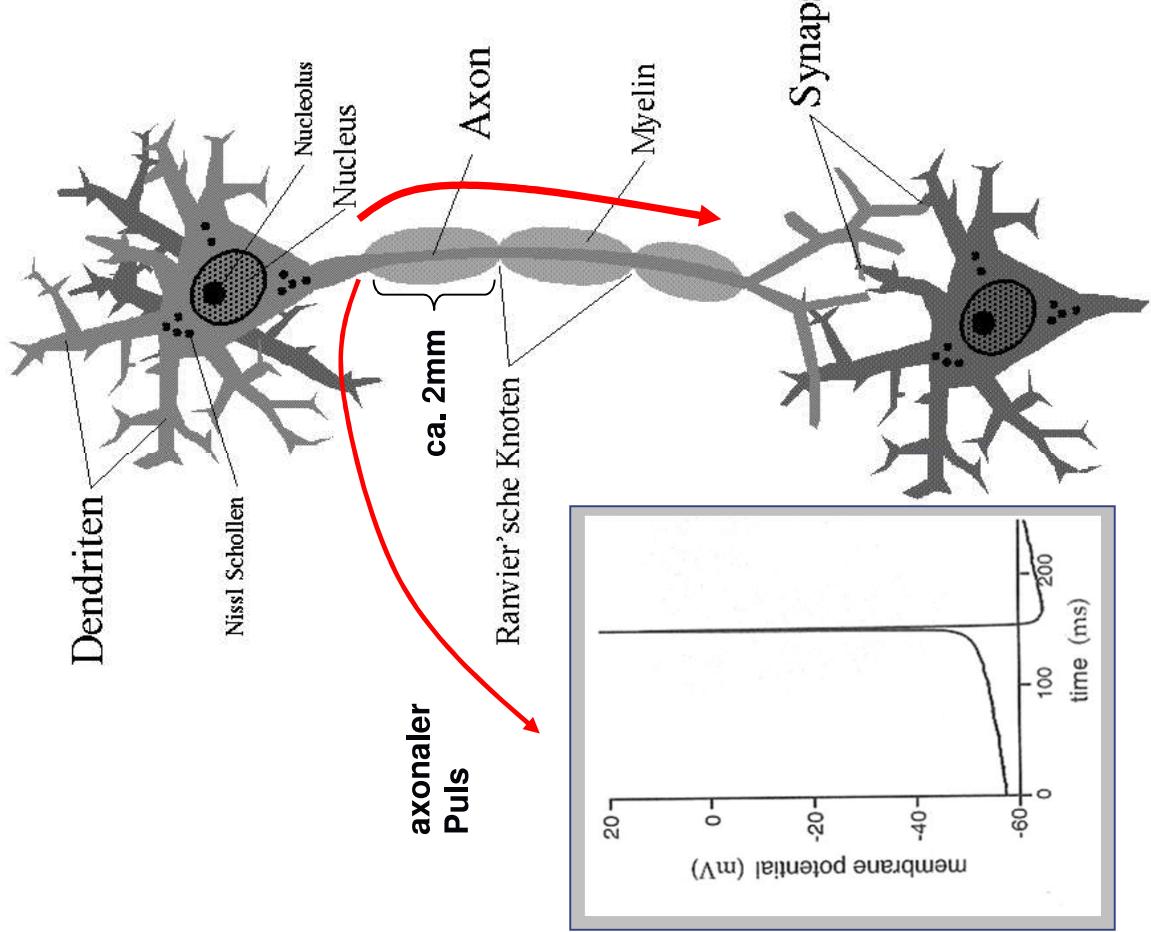
Begriff geprägt von Charles Scott Sherrington 1897

**Dendrit:** Fortsatz der Nervenzelle (griech. δένδρον *déndron*: Baum )  
Aufnahme synaptisch übertragener Nervenimpulse

**Axon, Neuraxon:** Fortsatz der Nervenzelle (griech. ἄξων *áxon* : Achse)  
Fortleitung elektrischer Nervenimpulse

**Axonhügel:** Ursprung des Axons, Ort der Generierung des Aktionspotentials  
(hohe Dichte spannungsaktivierter Natriumkanäle)





**Myelin** (griech. *myelos* = Mark, Gehirn)

Wird gebildet durch:

- + Schwann-Zellen (peripheres Nervensystem)
- + Oligodendrozyten (zentrales Nervensystem)

Strukturelle Beschreibung durch:  
Louis-Antoine Ranzier 1878

Wirkung: (I) Erhöhung des Übergangswiderstandes um den Faktor 5000  
(II) Erniedrigung der Kapazität um den Faktor 50

Ranvier'sche Knoten (Schnürringe):

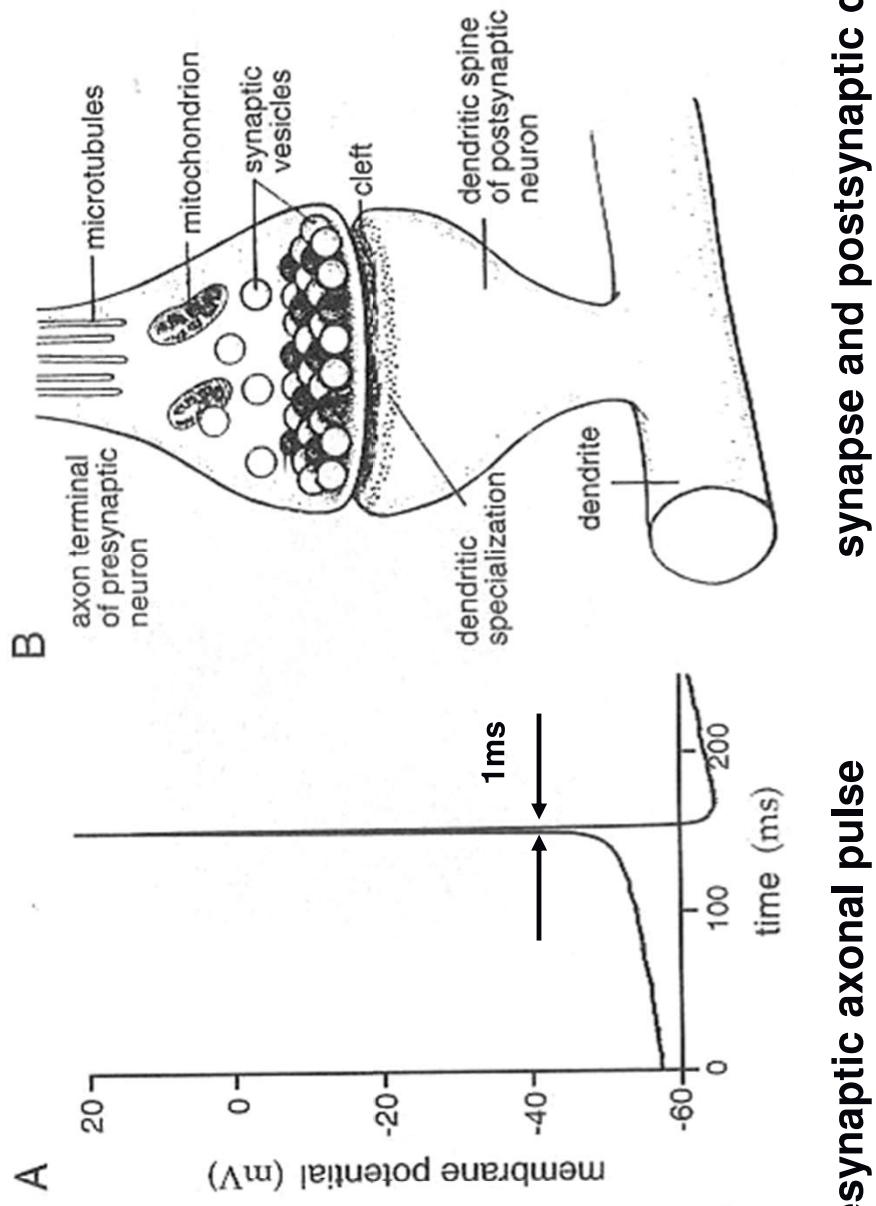
Ort der Entstehung von (regenerativen) Aktionspotentialen

Vorteile: (I) geringer Energieverbrauch

(II) schnellere Pulsortleitung

+ bis 180m/s (typ. 120m/s)

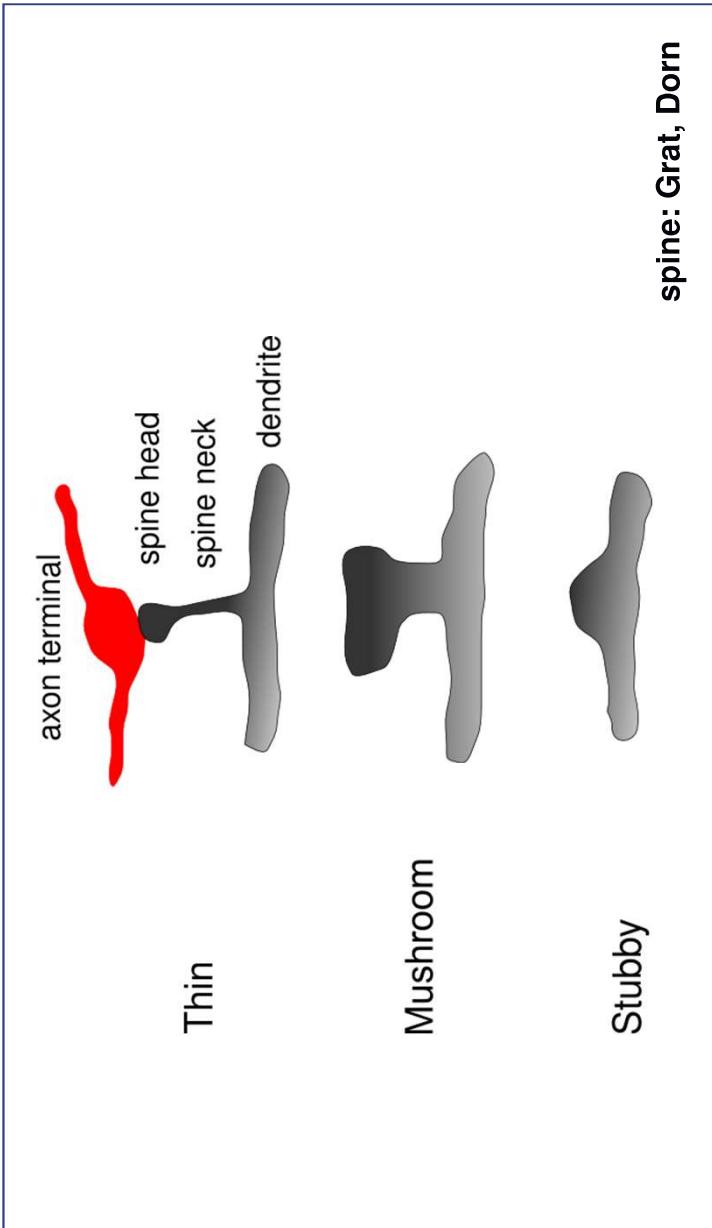




**presynaptic axonal pulse**

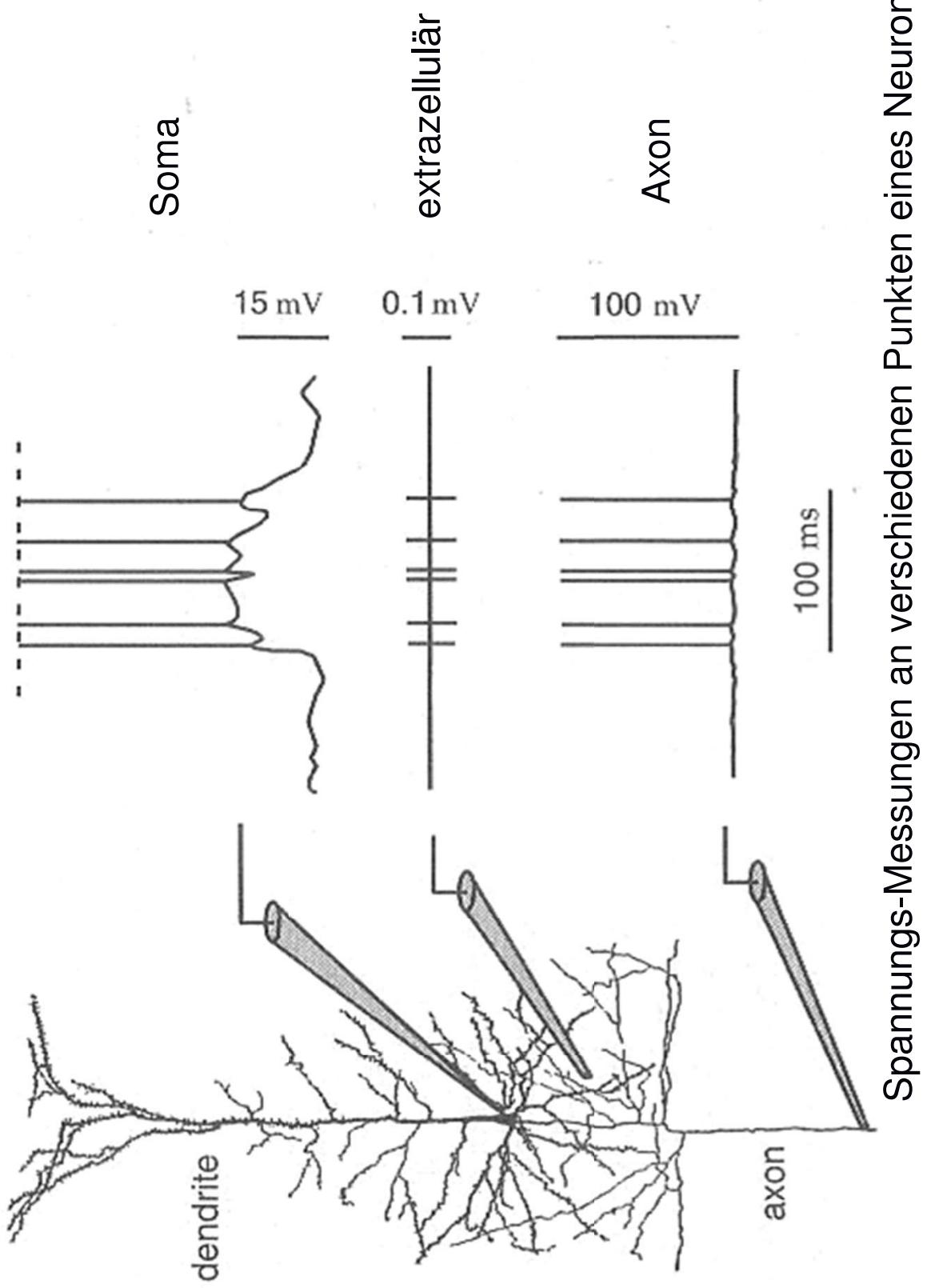
**synapse and postsynaptic dendrite**





- Zuerst entdeckt von Ramón y Cajal, 1888
- Funktion: Vergrößerung der Oberfläche des Dendriten
- Vergleich

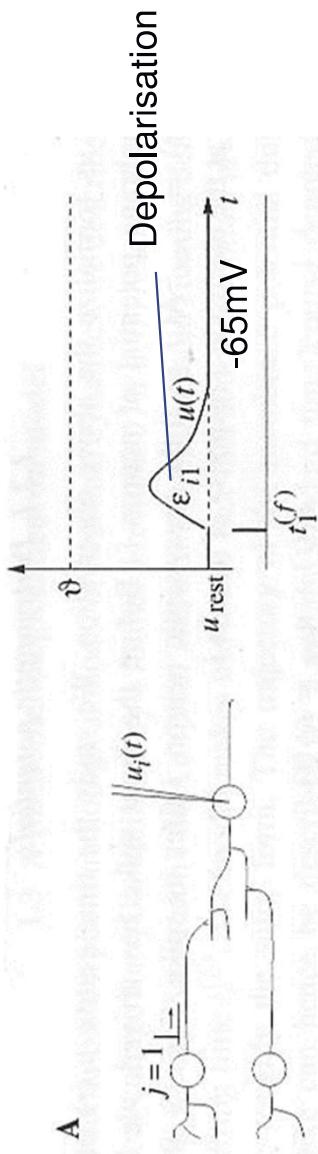
Ratte hat ca. 3000 Spines pro „apical dendrite“, Mensch hat ca. 6000 Spines pro Dendrit



Spannungs-Messungen an verschiedenen Punkten eines Neurons  
aus „Theoretical Neuroscience“

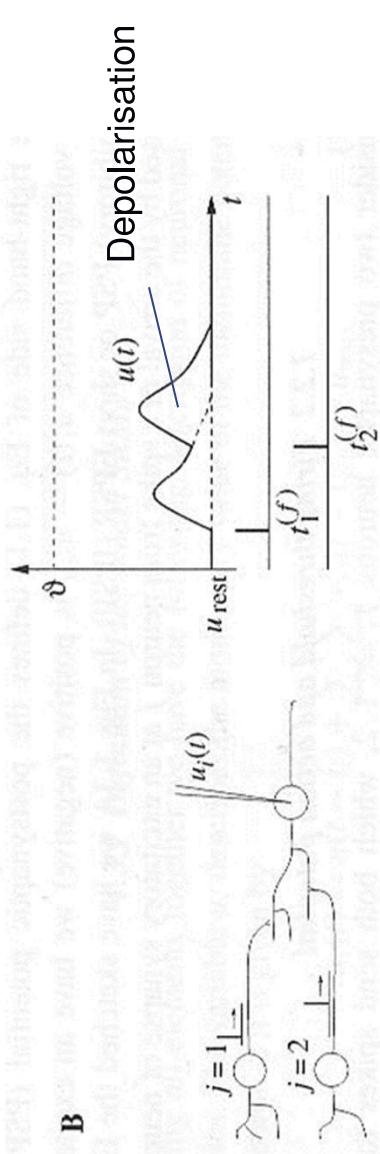
### A: einzelner Spike

Erzeugung eines EPSP  
(excitatory postsynaptic potential)

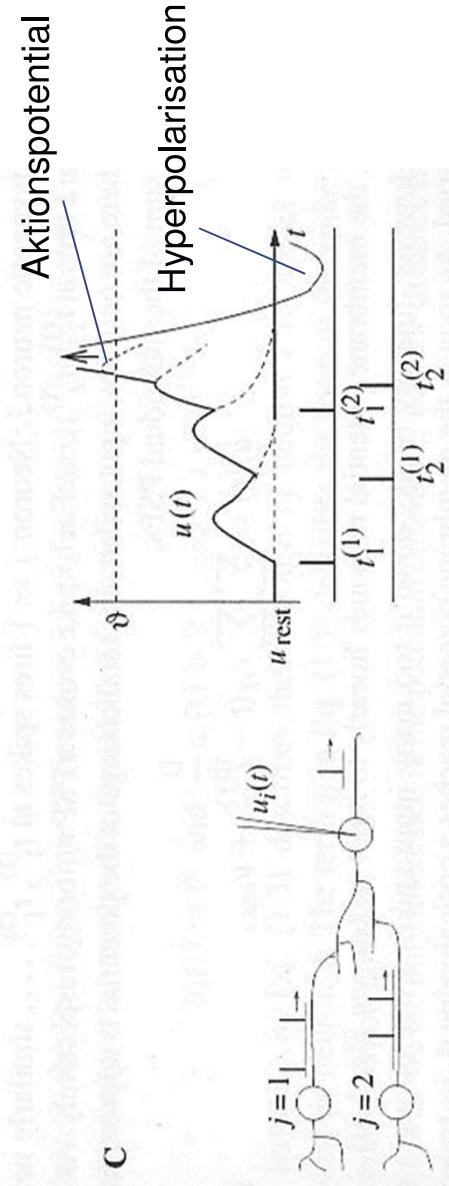


### B: zeitlich versetzte Spikes

Überlagerung von EPSPs



### C: Membranpotential überschreitet Schwelle



(aus „Spiking Neuron Models“)

- Erstmals mikrodynamische Beschreibung der Neuronen-Dynamik durch Hodgkin und Huxley (1952) -> „Giant Axon of the Squid“
  - Ergebnis: Aktions-Potentiale resultieren aus Ionen-Strömen, die durch die Zell-Membran fließen.
- 
- **Transportmechanismen für Ionen**  
Driftstrom, Diffusionsstrom, Ionen-Pumpen
  - **Ionen-Typen**  
 $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Ca}^{2+}$
  - **Transportmedium in der Zell-Membran**  
selektive Ionenkanäle mit spannungsabhängiger Dynamik

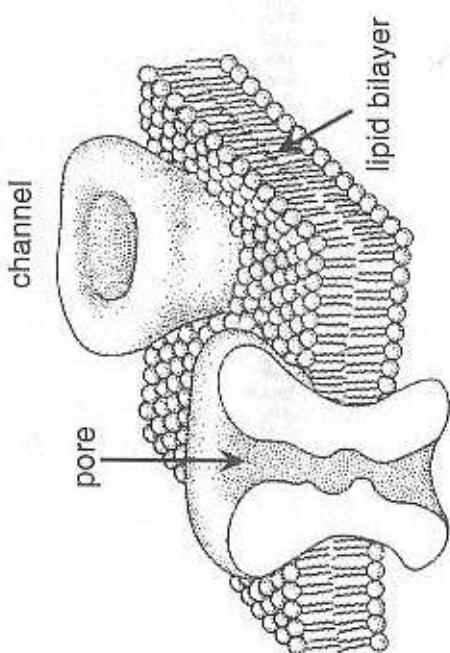
-> **Was ist zu modellieren?**

- Konzentrationsgradienten der Ionen (Diffusionsstrom)
- elektrische Feldstärkeverlauf über der Zellmembran (Driftstrom,  $u$ )
- Zeitkonstanten und Leitwerte für Ionenkanäle  $\tau(u)$   $g(u)$



## Bestandteile

- **Lipid-Doppelschicht**



vollkommen undurchlässig  
für Ionen und geladene Moleküle

- **Selektive Ionen-Kanäle**

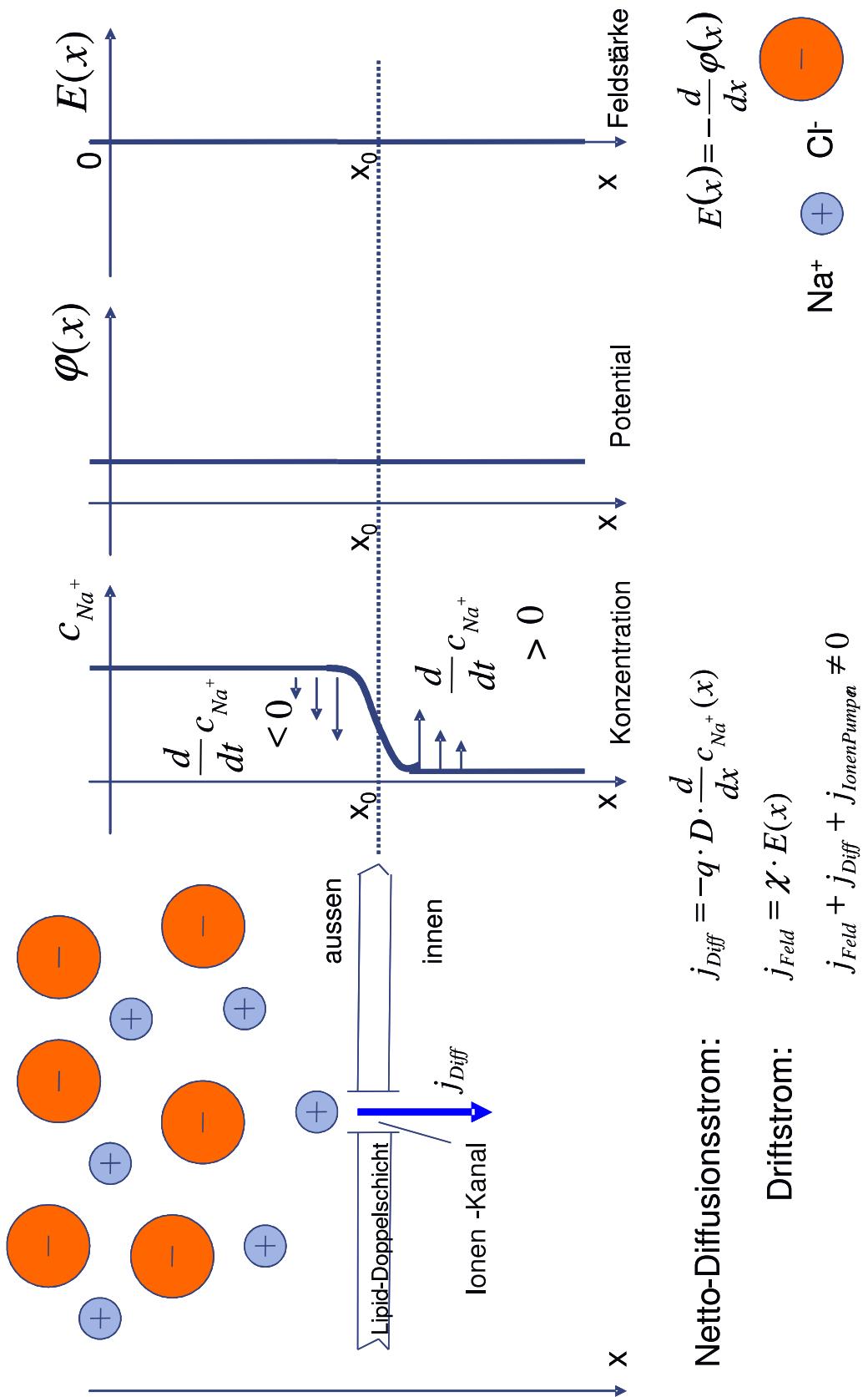


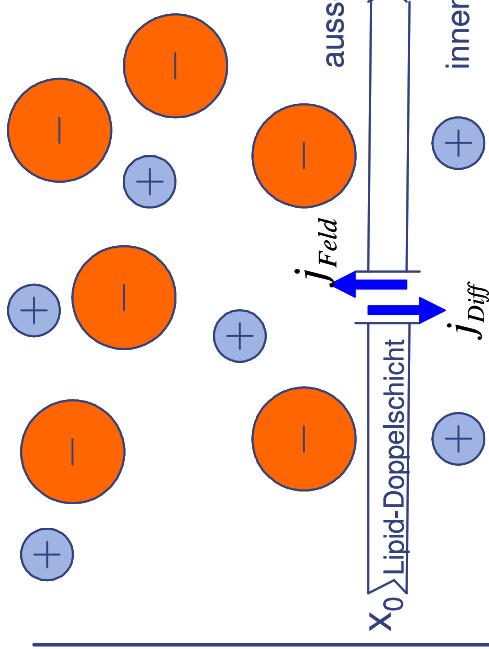
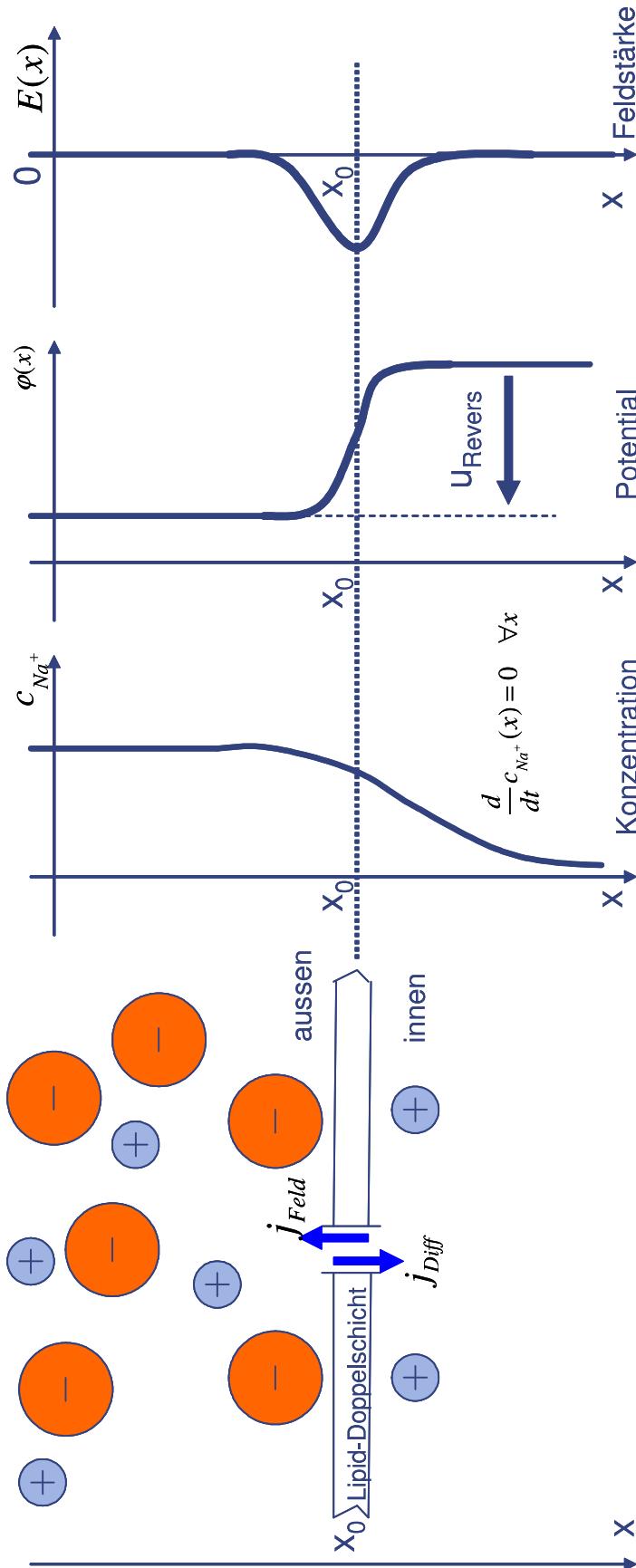
- **Ionenpumpen**

Schichtdicke der Membran: 3-4 nm  
Länge der Ionenkanäle: 10 nm

aus „Theoretical Neuroscience“

- (1) elektrogene Natrium-Kalium Pumpe
- (2) Natrium-Calcium Austauscher
- (3) Calcium-Pumpe





$$E(x) = -\frac{d}{dx} \varphi(x)$$

$$j_{Diff} = -q \cdot D \cdot \frac{d}{dx} c_{Na^+}(x)$$

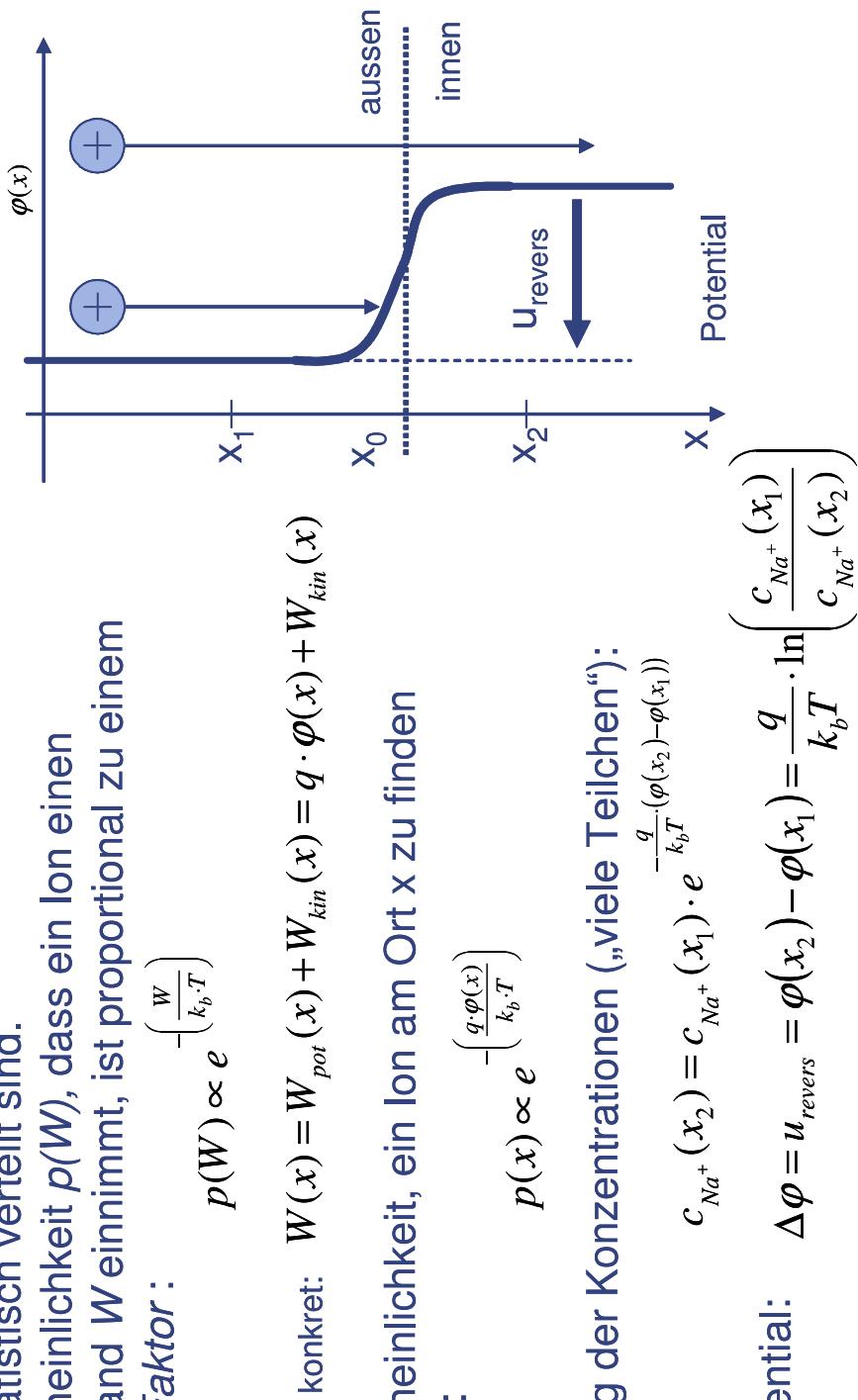
$$j_{Feld} = \chi \cdot E(x)$$

$$j_{Feld} + j_{Diff} + j_{IonenPumpen} = 0$$



- Potential stellt eine *Barriere* für die Diffusion geladener Teilchen dar.
- Diffusion ins Zell-Innere erfolgt nur dann, wenn die *kinetische Energie* ausreicht, die Barriere zu überwinden.
- Aus der Thermodynamik folgt, dass die *Energiezustände* der Ionen statistisch verteilt sind.

Die Wahrscheinlichkeit  $p(W)$ , dass ein Ion einen Energiezustand  $W$  einnimmt, ist proportional zu einem *Boltzmann Faktor*:

$$p(W) \propto e^{-\left(\frac{W}{k_b T}\right)}$$


- Die Wahrscheinlichkeit, ein Ion am Ort  $x$  zu finden ist demnach:

$$p(x) \propto e^{-\left(\frac{q \cdot \phi(x)}{k_b T}\right)}$$

-> Gewichtung der Konzentrationen („viele Teilchen“):

$$c_{Na^+}(x_2) = c_{Na^+}(x_1) \cdot e^{-\frac{q}{k_b T} \cdot (\phi(x_2) - \phi(x_1))}$$

-> Nernst Potential:

$$\Delta\phi = u_{revers} = \phi(x_2) - \phi(x_1) = \frac{q}{k_b T} \cdot \ln \left( \frac{c_{Na^+}(x_1)}{c_{Na^+}(x_2)} \right)$$

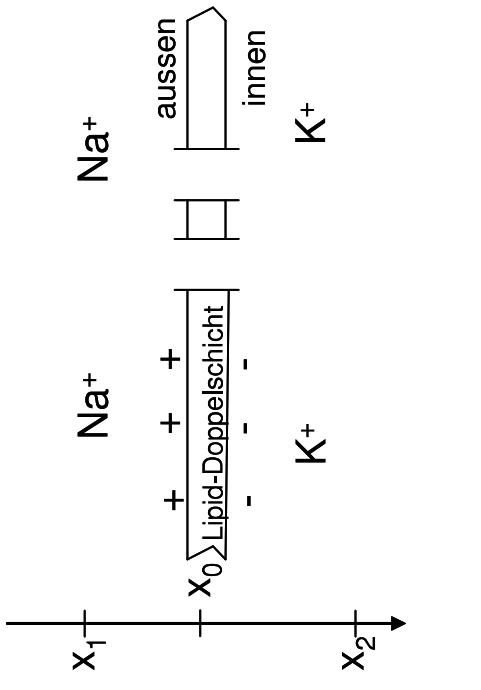
	$c(x_1) [mmol/l]$	$c(x_2) [mmol/l]$	$\Delta\varphi [mV]$
Na <sup>+</sup>	145	12	+50
K <sup>+</sup>	4	155	-77
Cl <sup>-</sup>	120	4	-60 ... -65
HCO <sub>3</sub> <sup>-</sup>	27	8	
A <sup>-</sup>	0	155	
Ca <sup>2+</sup>	0.4	—	150

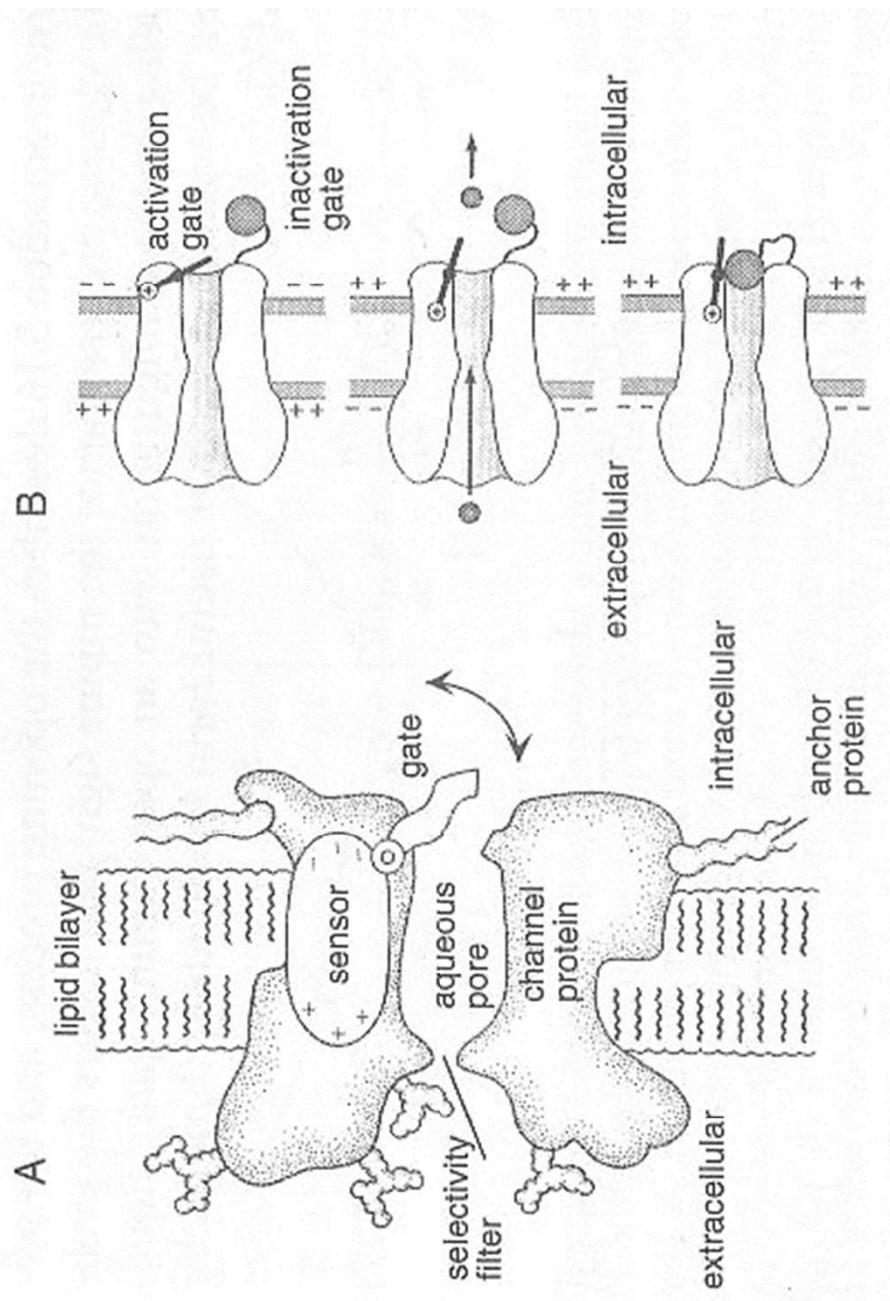
A<sup>-</sup>: Anionen

**Gesucht:** Gleichgewichtszustand des Systems  
**Problem:** Das Nernst Potential gilt nur für das *thermodynamisches Gleichgewicht*

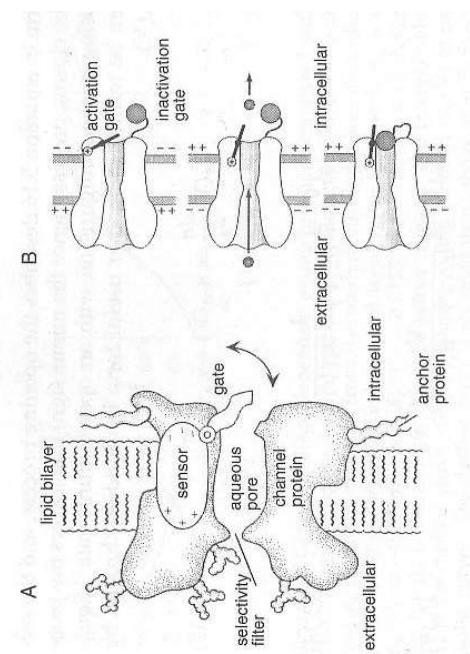
**Real:** • effektiver Feldstärkeverlauf E(x), el. Potential  
     • Nicht-Gleichgewichtskonzentrationen der Ionen  
     • stationäre Nettoströme durch die  
       verschiedenen Kanäle

Modellierung der dynamischen  
 Kanaleigenschaften erforderlich



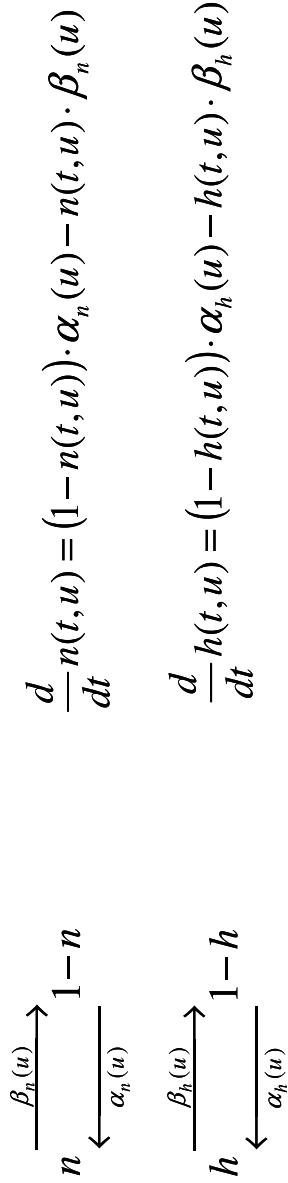


- Kanäle können in einem **offenen** (leitfähigen) oder **geschlossenen** (nicht-leitfähigen) Zustand sein.
- Kanäle wechseln ihren Zustand aufgrund einer Statistik, der Zustand eines Kanals wird mit sog. *Gating-Variablen* beschrieben, die fiktive *Gating-Partikel*/ modellieren.
  - *Gating-Partikel* kommen in zwei Zuständen
    - (i) **geschlossen** oder (ii) **offen** sowie zwei Arten
      - (a) **aktivierend** oder (b) **inaktivierend** vor.
  - Der *relative Anteil* offener aktivierender Partikel wird durch die Gating-Variablen  $n$  modelliert, analog für offene inaktivierende Partikel durch  $h$



aus „*Theoretical Neuroscience*“

- der Übergang eines Partikels von einem **offenen** Zustand zu einem **geschlossenen** Zustand und zurück wird durch *spannungsabhängige Raten* beschrieben:



## Umformulierung der Partikeldynamik in „Normalform“

$$\tau_n \cdot \frac{d}{dt} n(t) = -n(t) + n_\infty(u)$$
$$\tau_n = \frac{1}{\alpha_n(u) + \beta_n(u)}$$

$$n_\infty = \frac{\alpha_n(u)}{\alpha_n(u) + \beta_n(u)}$$

$$\tau_h \cdot \frac{d}{dt} h(t) = -h(t) + h_\infty(u)$$
$$\tau_h = \frac{1}{\alpha_h(u) + \beta_h(u)}$$

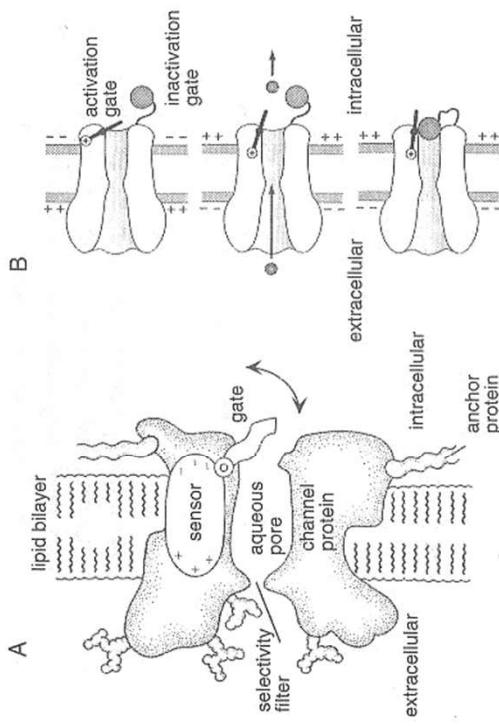
$$h_\infty = \frac{\alpha_h(u)}{\alpha_h(u) + \beta_h(u)}$$

**Zu beachten:** *Trägheit* und *Eregung* sind *spannungsabhängige Größen*

- Vorstellung: Kanäle bestehen aus einer Hintereinanderschaltung von  $q$  aktivierenden und  $p$  inaktivierenden Partikeln.
- Bedingung für Kanal-Leitfähigkeit: alle  $q$  aktivierenden und  $p$  inaktivierenden Partikel müssen offen sein.
- Hieraus folgt die Wahrscheinlichkeit, einen offenen Kanal vorzufinden (Partikel sind voneinander statistisch unabhängig):

$$P_i(t, u) = n^q(t, u) \cdot h^p(t, u)$$

$u$ : Membranspannung,  $t$ : Zeit,  $i$ : Index für Ionenotyp



- Man unterscheidet:

(A) **Persistente Kanäle**,  $p = 0$

(B) **Transient Kanäle**,  $p > 0$



- Modellierung der Zellmembran als Kapazität C
- Dynamik der Zellmembran:

$$I(t) = I_{\text{ionen}}(t) + C \cdot \frac{d}{dt} u(t)$$

- Betrachtung von drei verschiedenen Ionenströmen

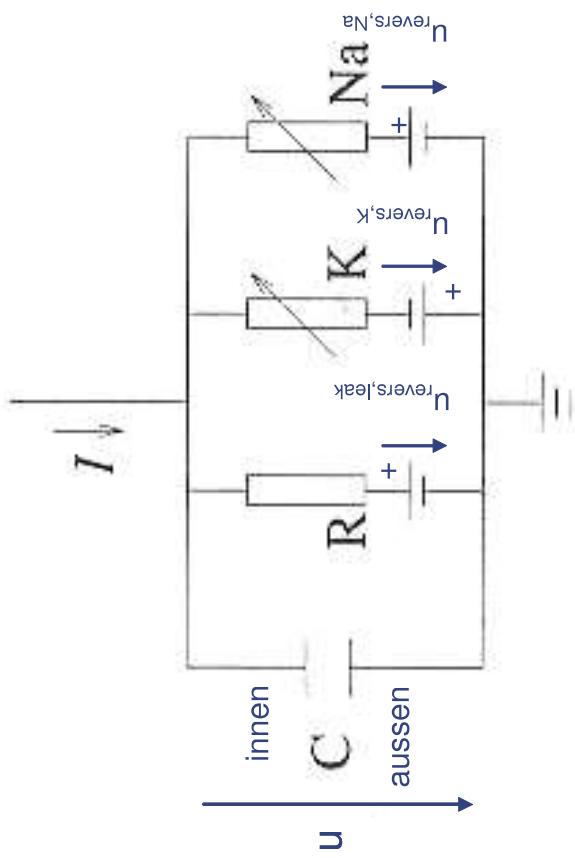
$$I_{\text{ionen}} = I_{Na} + I_K + I_{\text{Leak}}$$

- Abhängigkeit der Ionenstämme von Membranpotential und Reverspotentialen

$$I_i = G_i(u, t) \cdot (u(t) - u_{\text{revers}, i}) \quad i \in \{Na, K, \text{Leak}\}$$

- Leitwertmodellierung

$$G_i(t, u) = \bar{G}_i \cdot P_i(t, u) \quad i \in \{Na, K, \text{Leak}\}$$



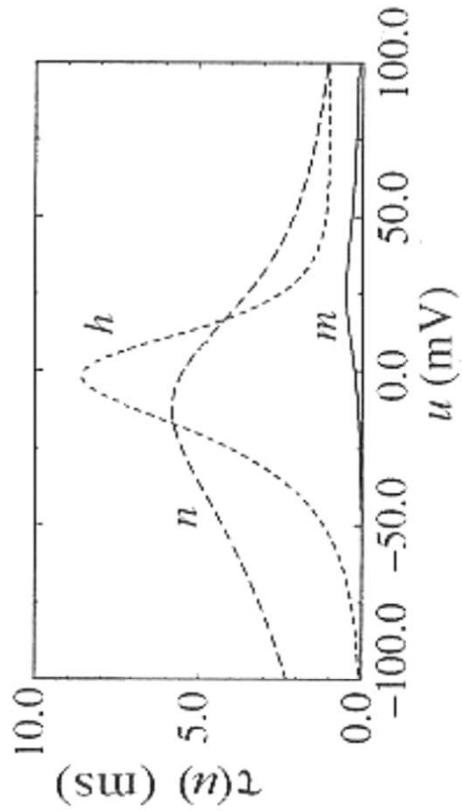
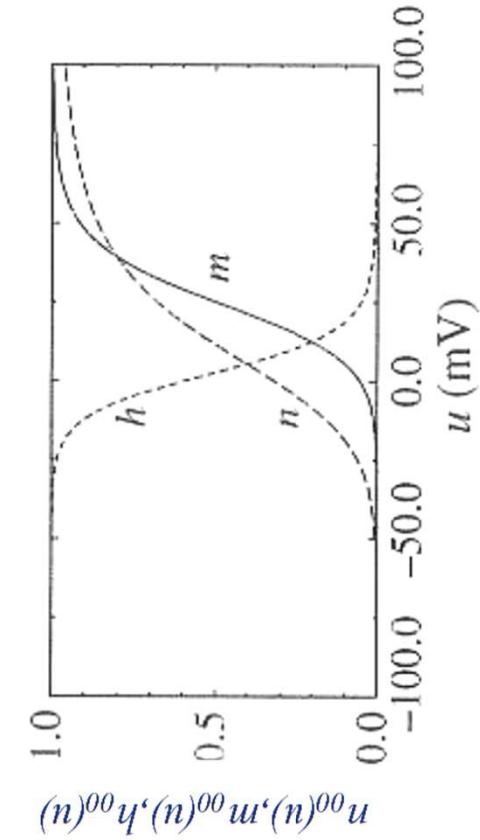
elektrisches Ersatzschaltbild der Zellmembran

$$\left. \begin{array}{l} G_{Na} = \bar{G}_{Na} \cdot m^3 \cdot h \\ G_K = \bar{G}_K \cdot n^4 \\ G_{\text{Leak}} = \bar{G}_{\text{Leak}} \end{array} \right\} \begin{array}{l} \text{m,n: aktivierende Partikel} \\ \text{h: inaktivierendes Partikel} \end{array}$$



	$\alpha_i(u[mV])$	$\beta_i(u[mV])$
n	$\frac{0.1 - 0.01 \cdot u}{e^{1-0.1 \cdot u} - 1}$	$0.125 \cdot e^{-\frac{u}{80}}$
m	$\frac{2.5 - 0.1 \cdot u}{e^{2.5-0.1 \cdot u} - 1}$	$4 \cdot e^{-\frac{u}{18}}$
h	$0.07 \cdot e^{-\frac{u}{20}}$	$\frac{1}{e^{3-0.1 \cdot u} + 1}$

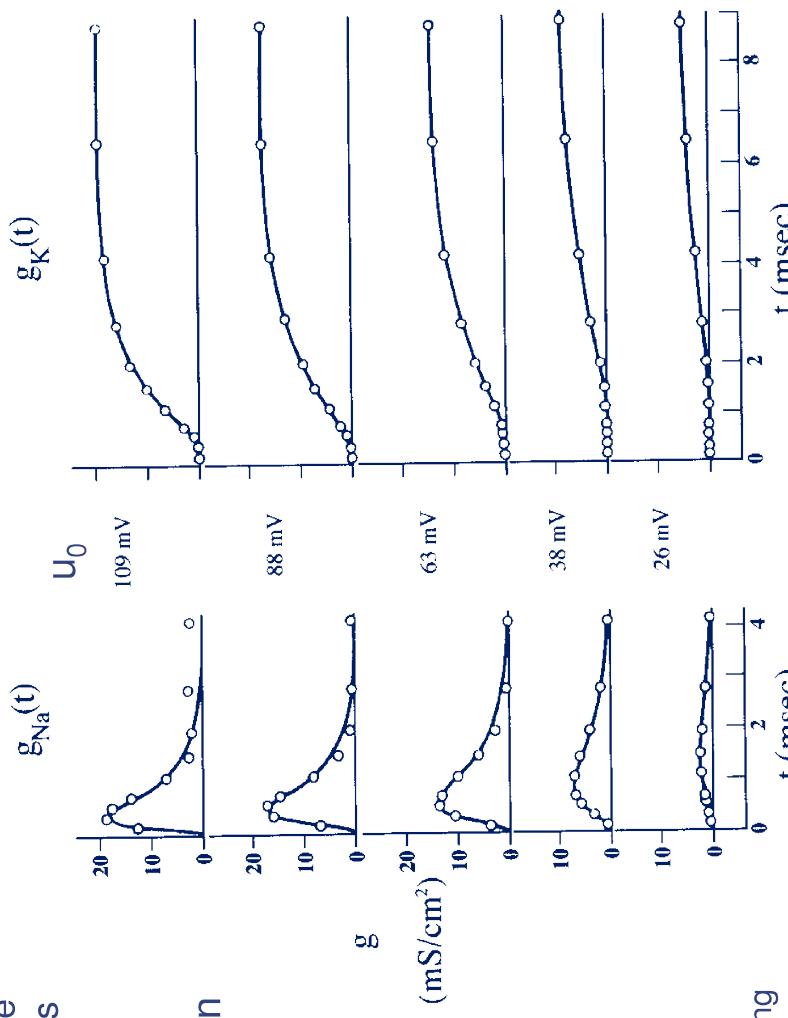
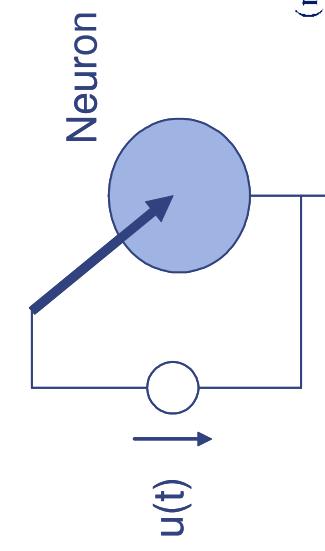
Anmerkung:  $u$  ist eine um  $U_{Rest}$  verschobene Spannung. Das Ruhpotential liegt hier bei  $u=0V$ .



(aus „Spiking Neuron Models“)

- Experiment: Erregung eines Neurons mit einem *Spannungsprung*

aufgrund der Selektivität der Kanäle  
kann so die zeitliche Dynamik eines  
Kanals bestimmt werden



$$u(t) = u_0 \cdot \chi(t) \quad \chi(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

- Beobachtungen

Na: erst Aktivierung, dann Inaktivierung

K: nur Aktivierung, langsam als Na-Kanal

— : Messung    ● : Theorie

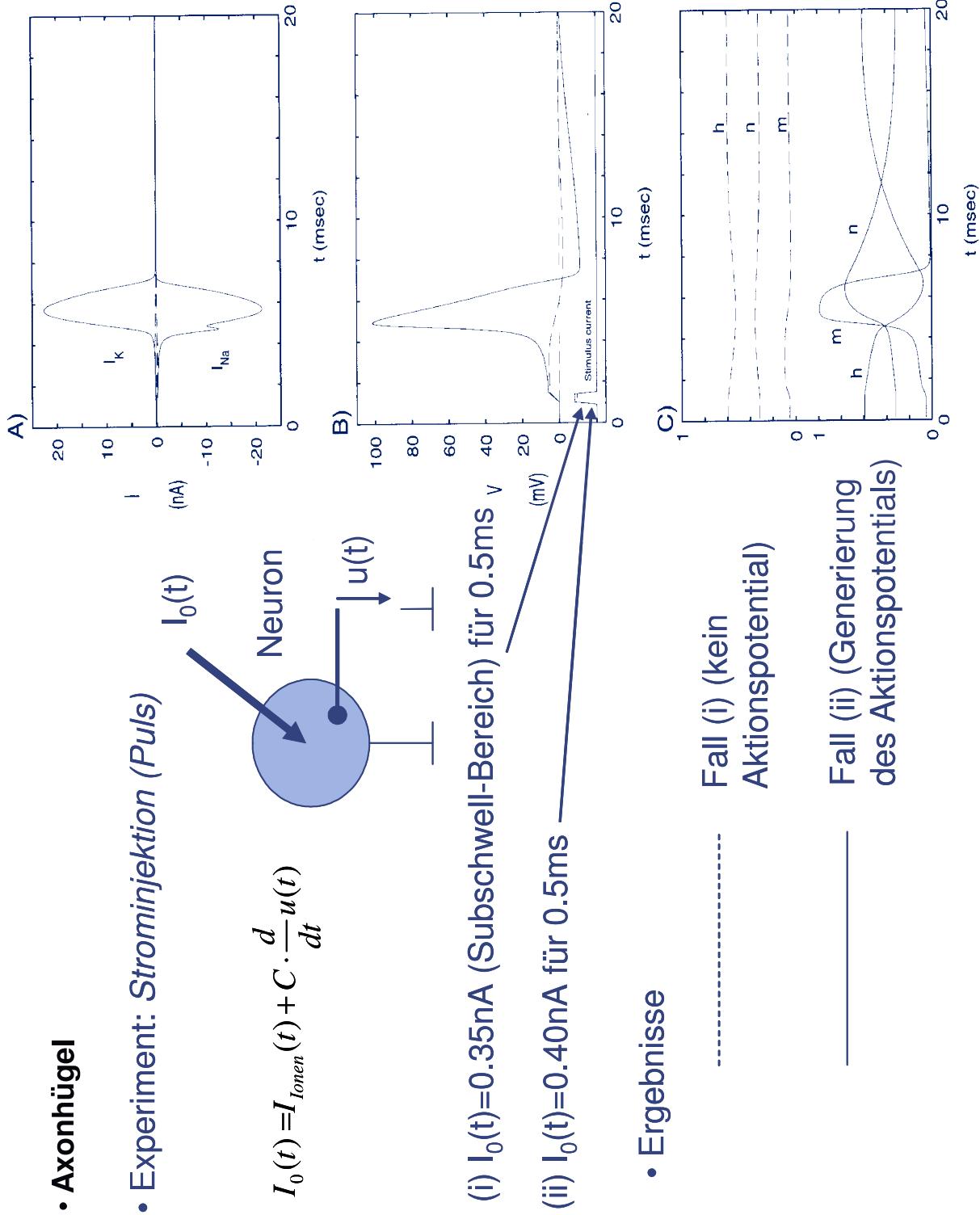
aus „Biophysics of Computation“

## Axonhügel

- Experiment: Strominjektion (Puls)

$$I_0(t) = I_{\text{ionen}}(t) + C \cdot \frac{d}{dt} u(t)$$

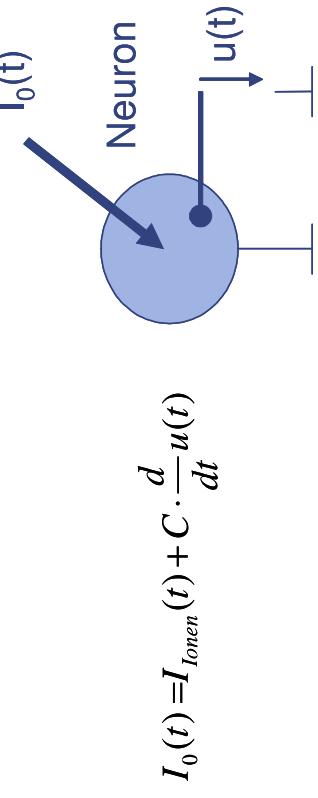
Neuron



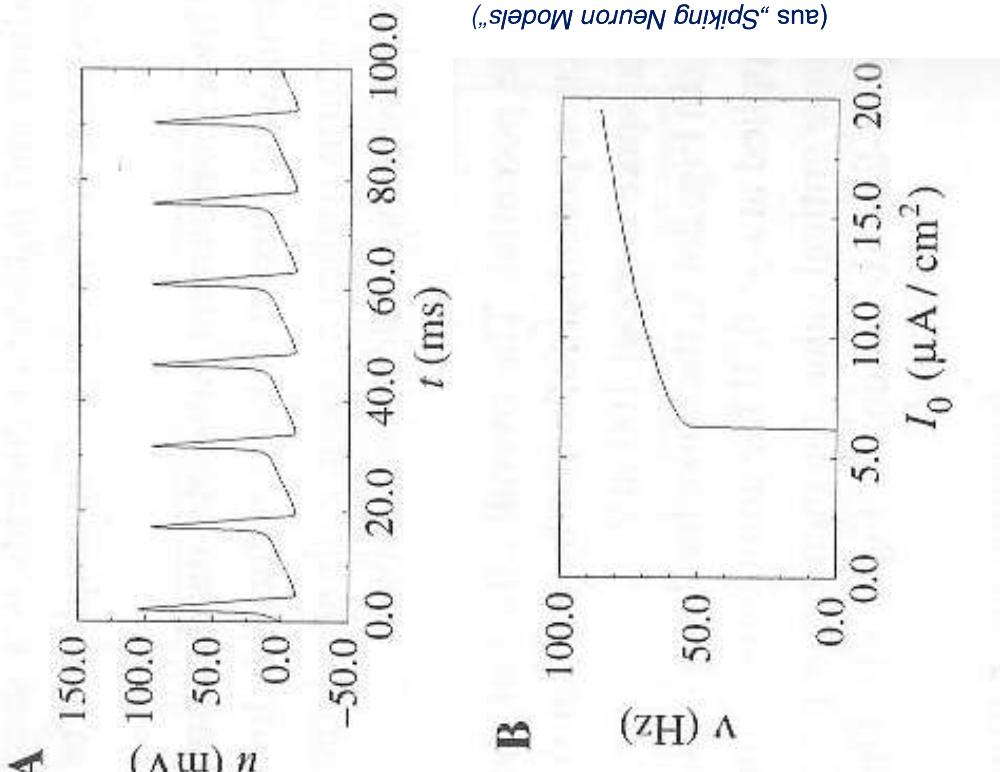
aus "Biophysics of Computation"



- Experiment: Strominjektion (konstant)



(aus „Spiking Neuron Models“)



### • Ergebnisse

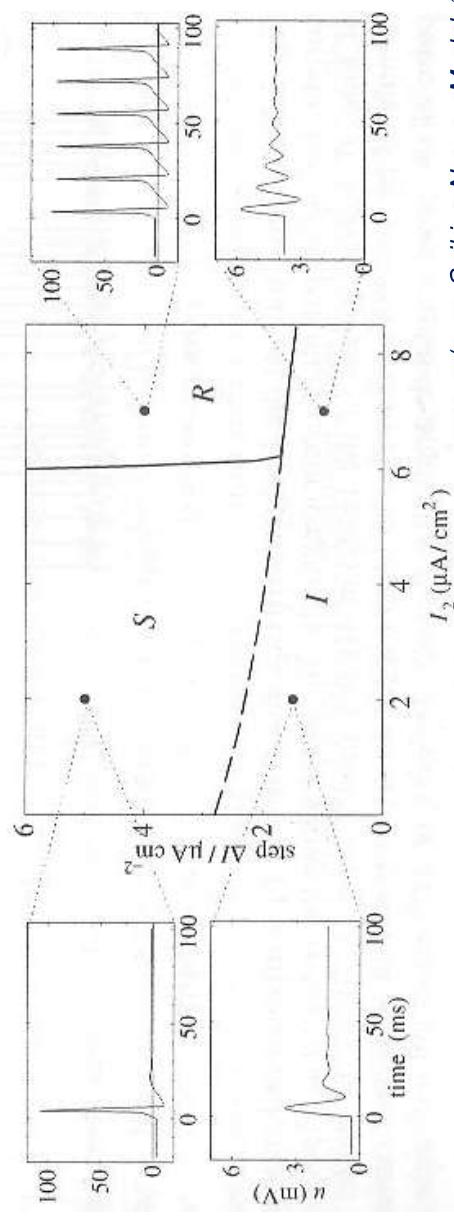
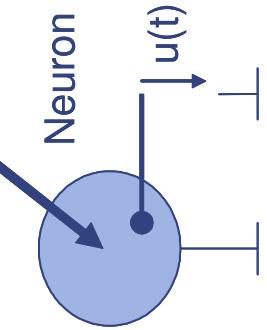
(A) Generierung eines Spike Train für hinreichend großes  $I_0$

$$I_\theta \approx 6 \mu\text{A}/\text{cm}^2$$

(B) Übertragungsfunktion (Pulsfrequenz aufgetragen über Strom  $I_0$ )  
-> Schwellenverhalten, Diskontinuität  
„Typ II“-Verhalten

- Experiment: Strominjektion (Sprung)

$$I_0(t) = I_1 + (I_2 - I_1) \cdot \chi(t) \quad I_0(t) = I_{\text{ionen}}(t) + C \cdot \frac{d}{dt} u(t)$$



- Ergebnisse

Es existieren Bereiche im  $I_1$ - $I_2$ -Raum, in denen

- keine Aktionspotentiale generiert werden (**Bereich I**)
- ein einzelnes Aktionspotential generiert wird (**Bereich S**)
- periodisch wiederkehrende Aktionspotentiale generiert werden (**Bereich R**)

(aus „Spiking Neuron Models“)

- Bisher bei Hodgkin-Huxley: Natrium-Kanal (transient), Kalium-Kanal (persistent) und Leckstrom
- Gültigkeit des Modells: z.B. „Giant Axon of the Squid“

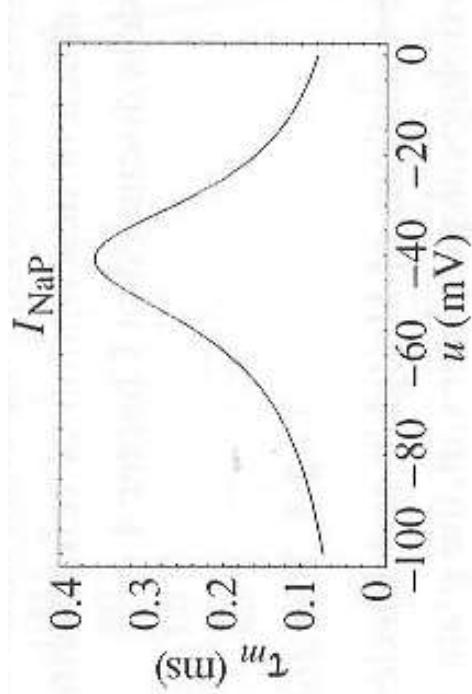
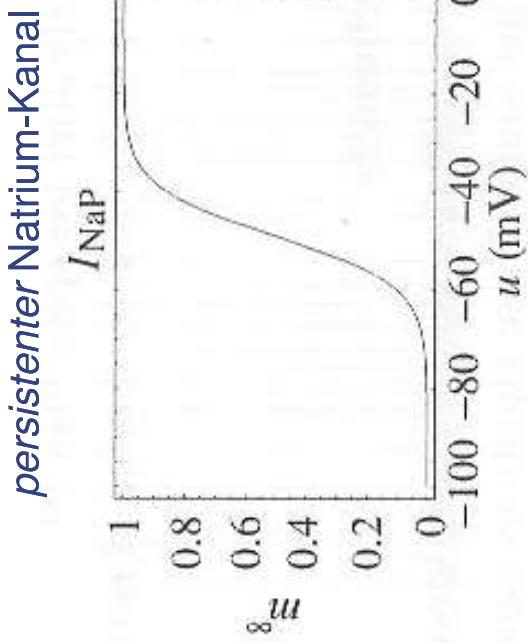
- Für viele Neuronen in höheren Organismen existieren eine weitaus größere Vielzahl an elektrophysiologischen Eigenschaften, die mit einer großen Variation verschiedener Kanäle einhergehen.

- allgemeines Modell für Kalium und Natrium:

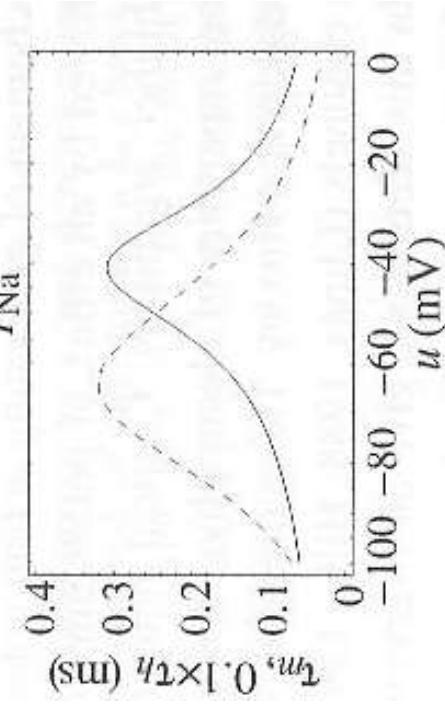
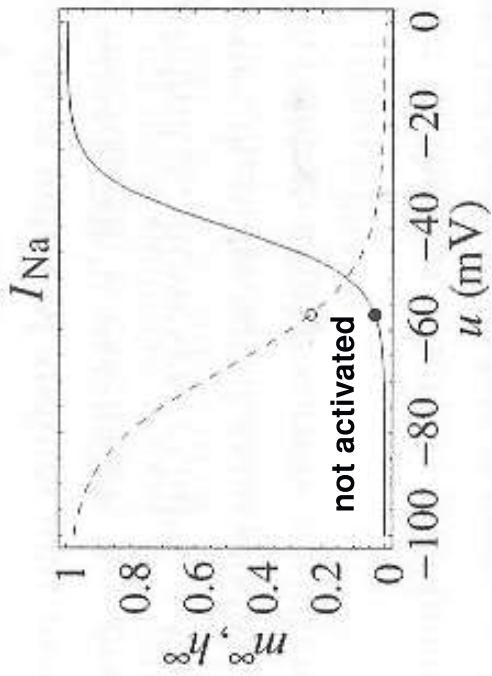
$$\text{Membran-Gleichung: } I_o(t) = I_{\text{ionen}}(t) + C \cdot \frac{d}{dt} u(t) \quad I_o(t): \text{ äußerer Erregungsstrom}$$

$$\text{Ionen-Kanäle: } I_{\text{ionen}}(t) = \sum_i I_i(t) \quad k : \text{Index für Kanaltyp}$$

$$\text{Ionen-Strom: } I_i(t) = \overline{G}_i \cdot n^{q_K}(t, u) \cdot h^{p_K}(t, u) \cdot (u - u_{\text{revers},i})$$



*transienter Natrium-Kanal*

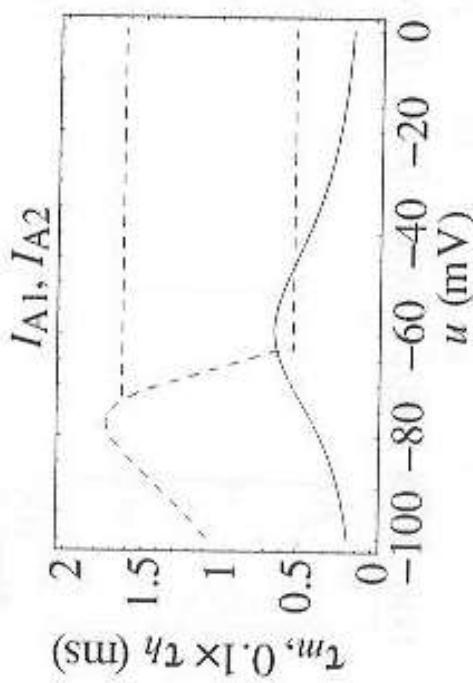
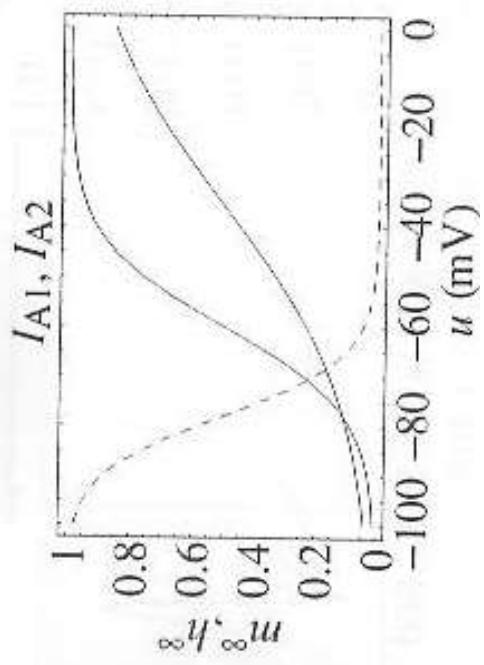


persistent

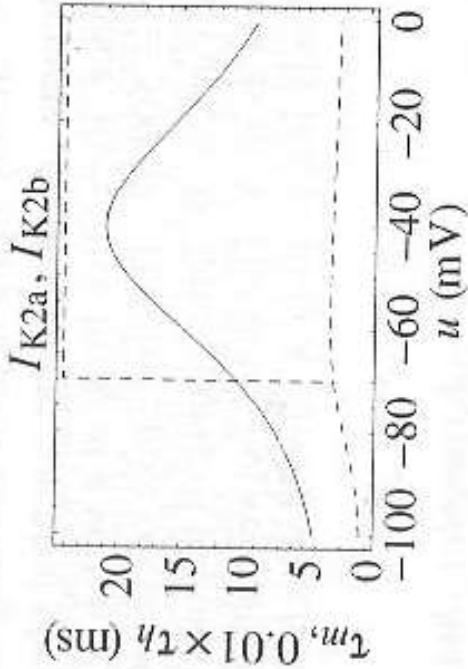
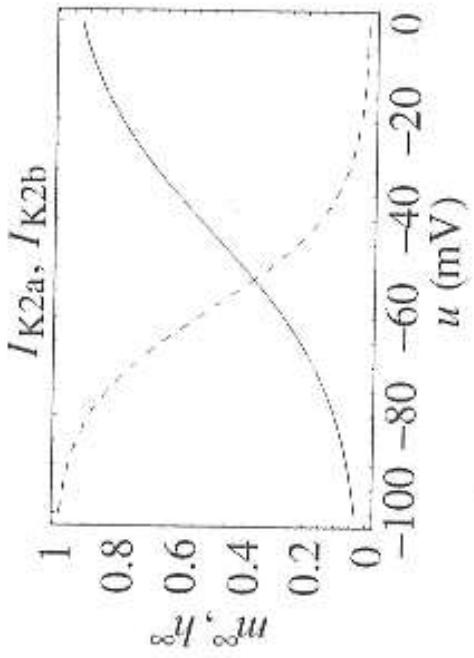
non-inactivating

erhöht die Erregbarkeit des Neurons  
durch Natrium-Einfluß im Falle der  
Depolarisation

## schnell-inaktivierender K-Kanal



## langsam-inaktivierender K-Kanal

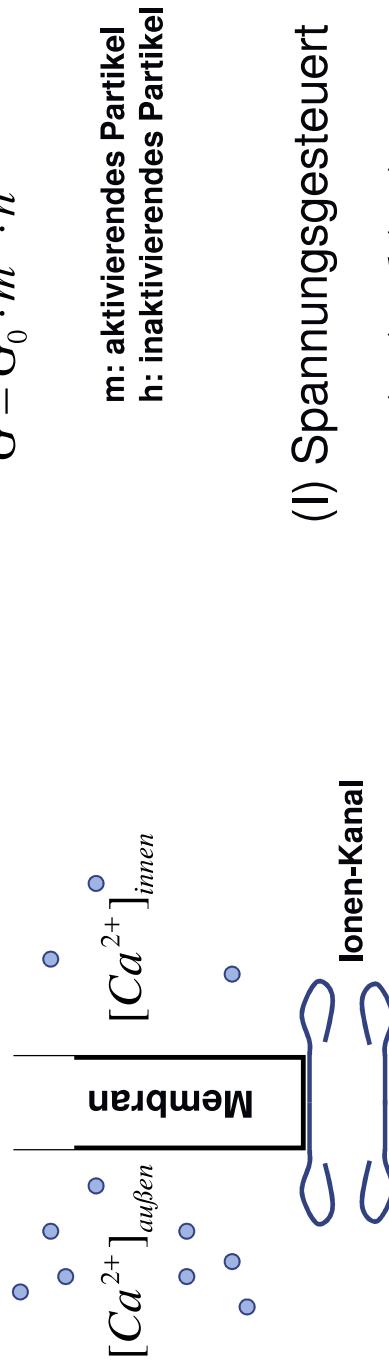


(aus "Spiking Neuron Models")



## Allgemeine Kanal-Leitwertformel

$$G = G_0 \cdot m^p \cdot h^q$$



(I) Spannungsgesteuert

$$m(u, t) \quad h(u, t)$$

(II) Calcium- und Spannungsgesteuert

$$m([Ca^{2+}]_{innen}, u, t) \quad h([Ca^{2+}]_{innen}, u, t)$$

(III) Calciumgesteuert

$$m([Ca^{2+}]_{innen}, t) \quad h([Ca^{2+}]_{innen}, t)$$

- Konzentration von Calcium in einer Zelle ist sehr gering und wird maßgeblich durch den Fluß durch die Ionenkanäle bestimmt.
- Aufbau und Abbau der inneren Ca-Konzentration:

$$\tau_{Ca} \cdot \frac{d}{dt} [Ca^{2+}]_i = -[Ca^{2+}]_i + \tau_{Ca} \cdot \phi_{Ca} \cdot I_X$$

- Beschreibung des Ionen-Flusses durch die *Goldman-Hodgkin-Katz-Gleichung*:

$$I_X = \bar{P}_X \cdot m_X^{g_X} \cdot h_X^{p_X} \cdot \frac{u \cdot z^2 \cdot F^2}{RT} \cdot \frac{[Ca^{2+}]_e \cdot e^{-uzF/RT} - [Ca^{2+}]_i}{1 - e^{-uzF/RT}} \quad z = 2 \text{ (Ca}^{2+}\text{)}$$

$[Ca^{2+}]_e$  : zell-externe Calcium-Ionen Konzentration (0.4 m mol<sup>-1</sup>)

$[Ca^{2+}]_i$  : zell-interne Calcium-Ionen Konzentration

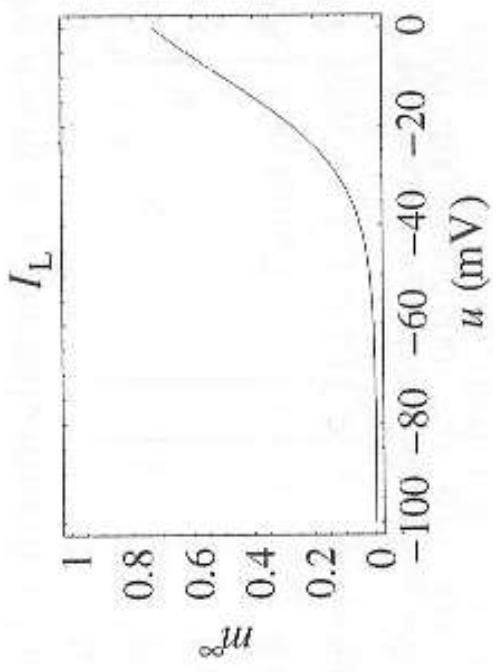
$\bar{P}_X$  : maximale Permeabilität

$$F = 9.648 \cdot 10^4 C mol^{-1} \text{ (Faraday Konstante)}$$

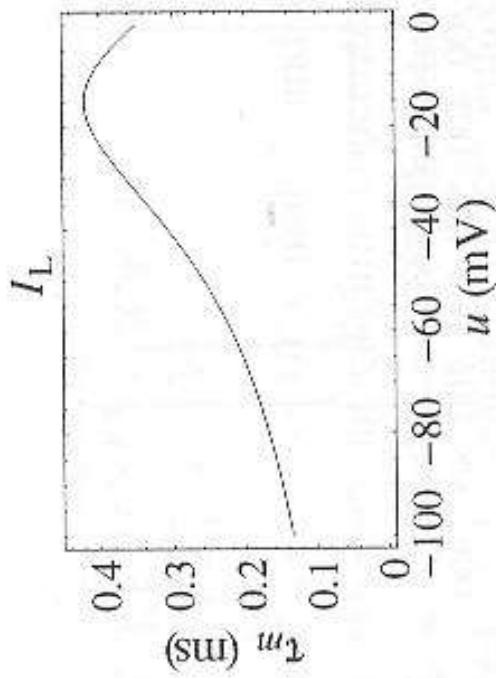
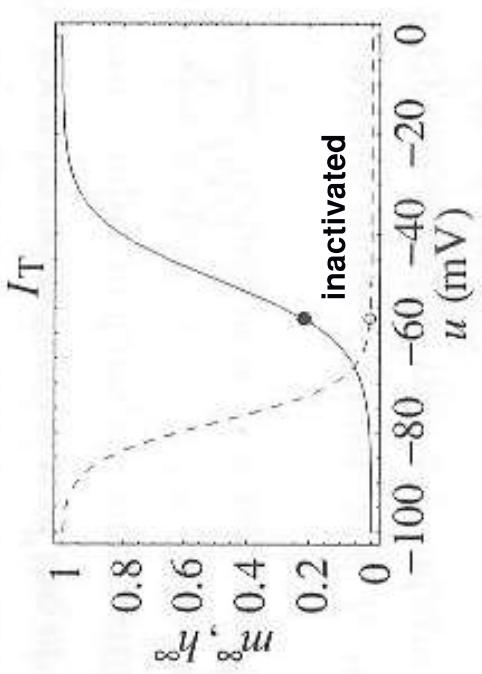
$$R = 8.314 V C K^{-1} \text{ (Gas Konstante)}$$



## High-Threshold Calcium-Kanal



## Low-Threshold Calcium-Kanal



(aus "Spiking Neuron Models")

- $I_C$  und  $|I_{AHP}|$
- C: Calcium-activated AHP: after hyperpolarization
- Modell des Kalium-Stroms:

$$I_{C,AHP} = \bar{G}_{C,AHP} \cdot m \cdot (u - u_{Revers,K})$$

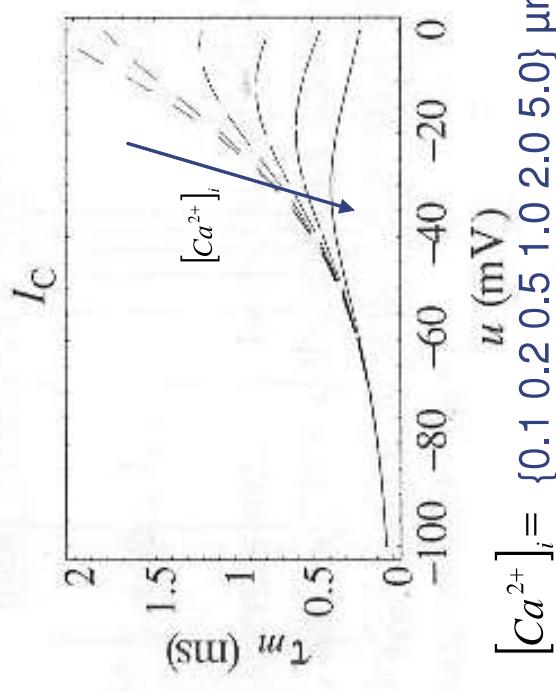
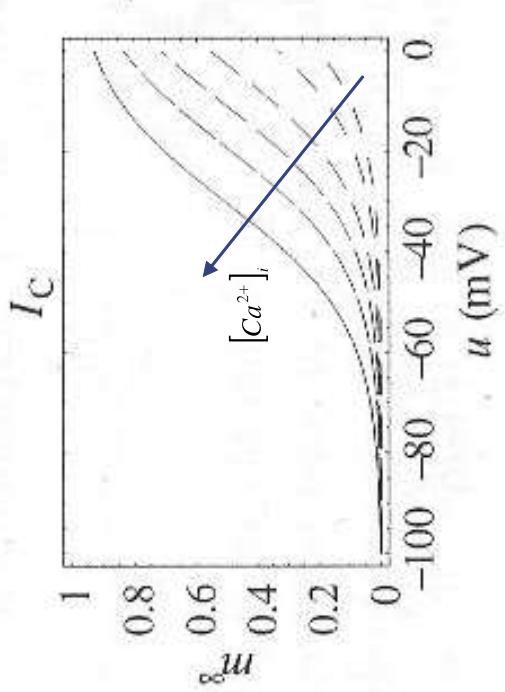
- Gleichung für aktivierendes Partikel:

$$\frac{d}{dt}m = \alpha \cdot m - \beta \cdot (1-m)$$

- Calciumkonzentrationsabhängige Übergangsraten:

$$\alpha = 2.5 \cdot 10^5 \cdot [Ca^{2+}]_i \cdot e^{\frac{u}{24}} \quad \beta = 0.1 \cdot e^{-\frac{u}{24}}$$

$$|I_{AHP}| = \min(c \cdot [Ca^{2+}]_i, 0.01) \quad \beta = 0.001$$



(aus „Spiking Neuron Models“)



## Refraktärzeit

Neuron kann nach Aussenden eines Spikes erst nach Ablauf der Refraktärzeit wieder zum Feuern angeregt werden

## Spikeratenadaptation

Bei konstantem Input steigen die zeitlichen Abstände zwischen den ausgangsseitigen Pulsen („Gewöhnung“)

## Postinhibitory Rebound

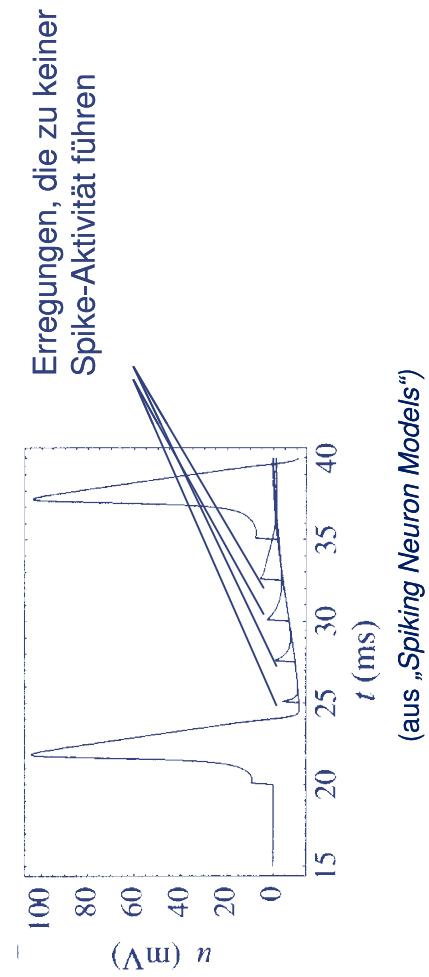
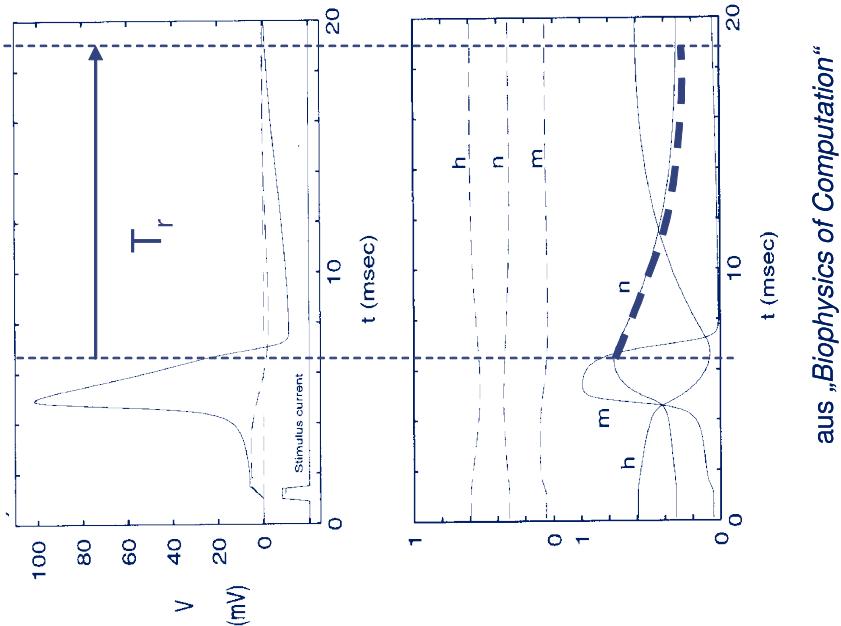
Beim Abschalten einer eingesetzten Inhibition sendet das Neuron einen oder mehrere Pulse aus

## Übertragungsfunktion Typ I und Typ II

Kontinuierliche Übertragungsfunktion (f-i-Diagramm)



- experimentelle Beobachtung: Neuronen lassen sich nach dem Aussenden eines Spikes erst nach einer sog. „Refraktärzeit“  $T_r$  wieder zum Feuern anregen
  - Ursache: Das n-Partikel (Kalium-Kanal) ist nach dem Spike noch signifikant aktiv und hyperpolarisiert die Membran
  - Strominjektionen können hier die Membran nur unzureichend depolarisieren



- Wirkung des  $I_{AHP}$ -Strom (Ca-aktivierter Kalium-Strom)

$$I_{AHP} = \bar{G}_{AHP} \cdot m \cdot (u - u_{revers,K})$$

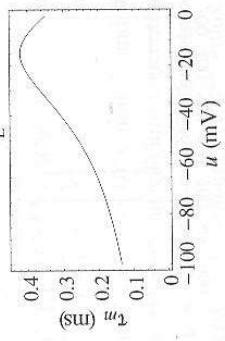
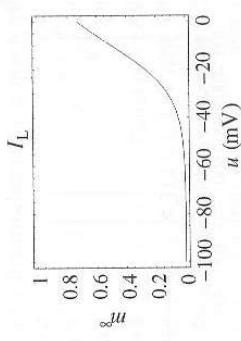
$$\frac{d}{dt} m = \alpha \cdot m - \beta \cdot (1-m)$$

$$\alpha = \min(c \cdot [Ca^{2+}]_i, 0.01)$$

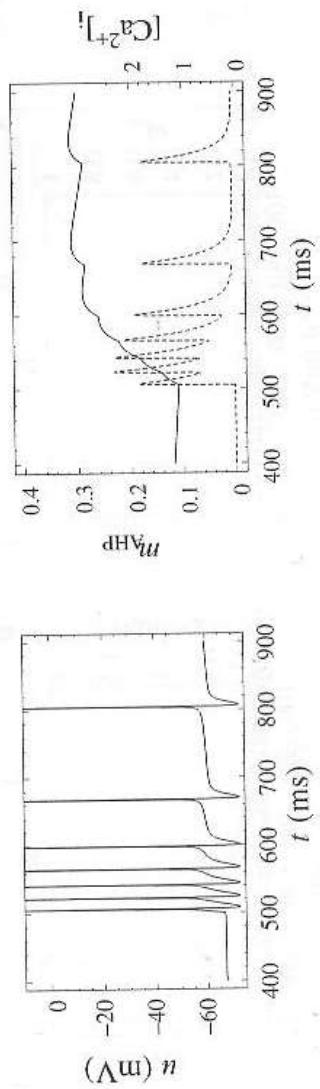
$$\beta = 0.001$$

- Calcium wird durch L-Kanal während eines Spikes in die Zelle transportiert. Folge: Mit jedem Spike steigt die zell-innere Ca-Konzentration, wodurch  $I_{AHP}$  aktiviert wird mit der Tendenz zur Hyperpolarisation

- Konsequenz: Depolarisation der Membran wird zunehmend durch  $I_{AHP}$  gehemmt -> Spikerate fällt

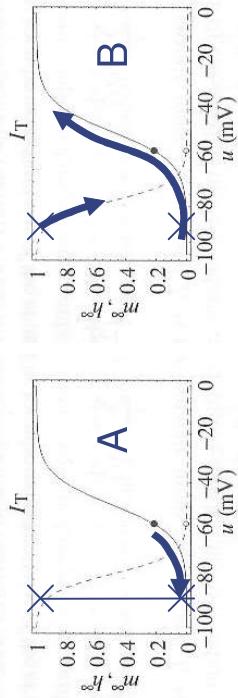


(aus „Spiking Neuron Models“)

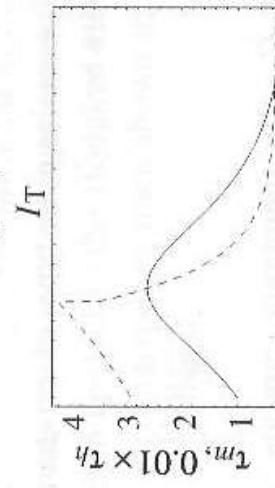


firing frequency adaptation

- Der Calcium T-Kanal hat inaktivierende Partikel mit hoher Zeitkonstante
- bei starker Hyperpolarisation (durch Injektion eines hyperpolarisierenden Stroms) der Membran wird der T-Strom zunächst inaktiv (A):

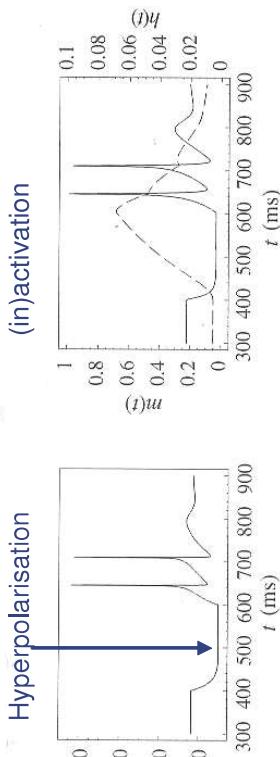


- Das Abschalten des hyperpolarisierenden Stroms bewirkt eine Depolarisation.  $I_T$  wird aktiviert (B) und lädt die Membran auf. Zusätzlich folgt der transiente Na-Strom.



- Folge: ein oder mehrere Spikes werden ausgelöst

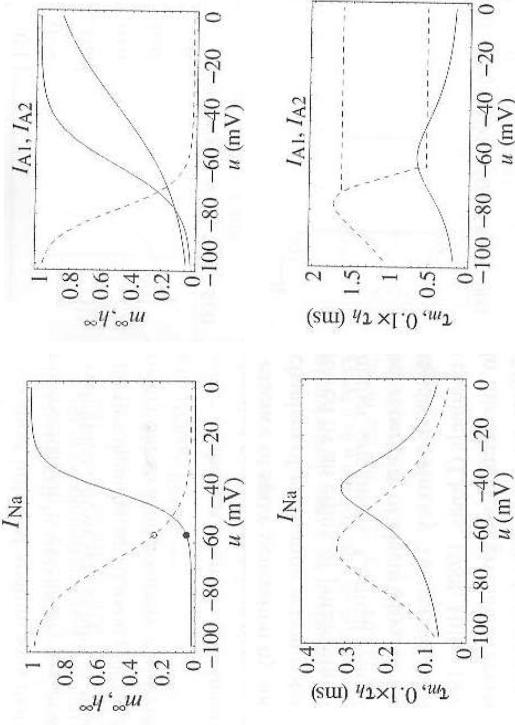
IPSP: inhibitory postsynaptic potential



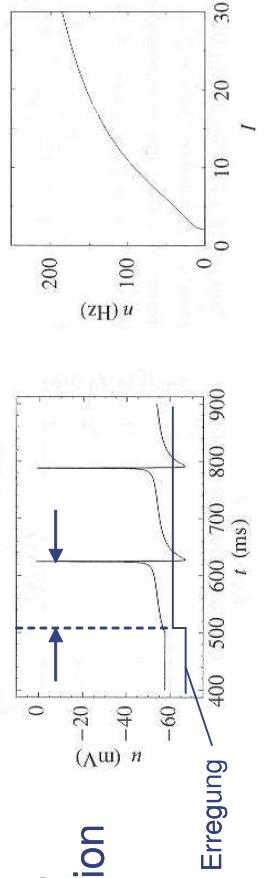
(aus „Spiking Neuron Models“)

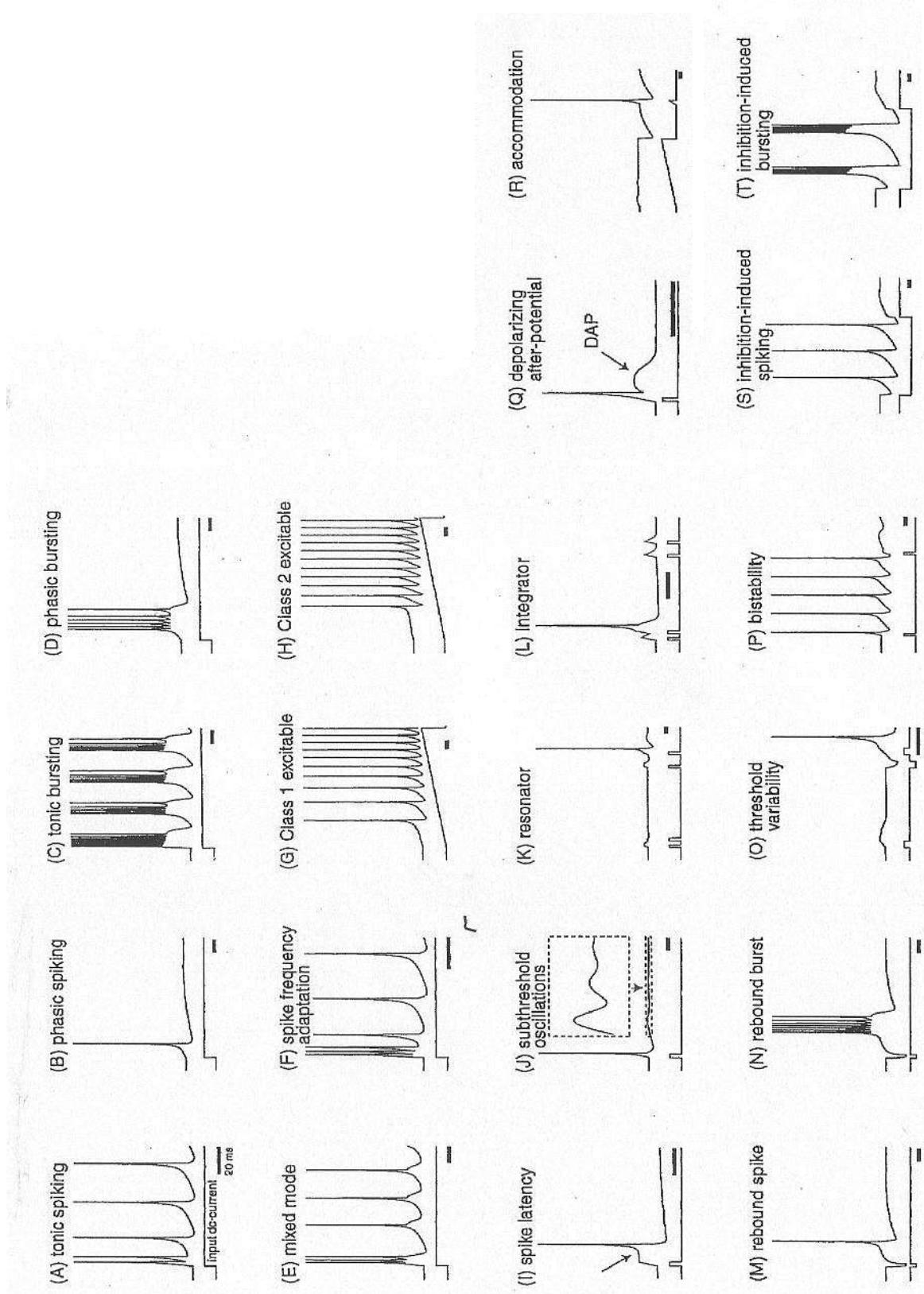


- Der **A-Strom** ist ein *transienter* Kalium-Strom
- Wirkung: Hyperpolarisierend, positive Ladung wird aus der Zelle transportiert.
- Zeitkonstanten der inaktivierenden Partikel des A-Stroms sind deutlich größer als bei dem transienten Natrium-Kanal, A-Strom wirkt lange hyperpolarisierend
- Für eine Depolarisation der Membran durch einen Natrium-Strom muss erst der Kalium-Strom inaktiviert sein
- Konsequenz:



(aus „Spiking Neuron Models“)





aus E.M. Izhikevich, "Which Model to Use for Cortical Spiking Neurons?", *IEEE Transactions on Neural Networks*, 2004

