# Algebraische Komplexitätstheorie III

## Zur Berechnungskomplexität von Permanenten

MICHAEL CLAUSEN, UNIVERSITÄT BONN

Im dritten und letzten Teil unserer kleinen Vortragsreihe über algebraische Komplexitätstheorie ging es um ein algebraisches Analogon zur Theorie der **NP**-Vollständigkeit. Dieses Analogon geht auf Valiant [15, 17] zurück und entsprang seinen Studien von Zählproblemen [14]. Neben den offensichtlichen Querverbindungen zur Kombinatorik werden durch dieses Thema aber auch innerhalb der Komplexitätstheorie Brücken geschlagen: einerseits eine Brücke zur strukturellen Komplexitätstheorie, bei der es (etwa auf der Grundlage des Turingmaschinenmodells) um die Formulierung von Komplexitätsklassen und deren Beziehungen untereinander geht, andererseits eine Brücke zur parallelen Komplexitätstheorie.

Diese Vortragsausarbeitung beginnt mit einer Erinnerung an die Booleschen Komplexitätsklassen **P** und **NP** sowie an den Begriff der **NP**-Vollständigkeit. Danach werden die algebraischen Analoga **VP** und **VNP** einführt sowie der Begriff der **VNP**-Vollständigkeit vorgestellt. Während die Objekte im Booleschen Fall Sprachen sind, also Mengen von Wörtern endlicher Länge über einem endlichen Alphabet, sind die Objekte im algebraischen Fall gewisse unendliche Folgen multivariater Polynome über einem Körper $k$. Wichtige Rollen im algebraischen Analogon werden die Folgen $DET = (DET_n)$ bzw. $PER = (PER_n)$ der generischen Determinanten bzw. Permanenten spielen:

$$DET_n := \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} X_{i\sigma(i)}\ , \quad PER_n := \sum_{\sigma\in S_n} \prod_{i=1}^{n} X_{i\sigma(i)}.$$

Wir werden die (erweiterte) Hypothese von Valiant diskutieren, aus deren Gültigkeit sich ergeben würde, daß es zwischen der Berechnungskomplexität von Determinanten und Permanenten trotz der Ähnlichkeit in der Definition krasse Unterschiede gibt. Obwohl viele Indizien für diese Hypothese sprechen, ist man von einem Beweis noch sehr weit entfernt.

# 1 Die Hypothesen von Cook und Valiant

In der ersten Hälfte dieses Jahrhunderts hat man die Frage nach dem prinzipiell Berechenbaren auf verschiedene, aber äquivalente Weise befriedigend formalisieren können durch Konzepte wie Turingmaschinen, Registermaschinen, WHILE-Programme, rekursive Funktionen, Thue-Systeme und Markov-Algorithmen. Zu Beginn des Computer-Zeitalters trat dann naturgemäß die Frage in den Vordergrund: *Welche Probleme sind in einem praktischen Sinne berechenbar?*

Als erste Approximation der in einem praktischen Sinn berechenbaren Probleme sieht man die Komplexitätsklasse **P** an, die aus allen Sprachen $A$ über einem Alphabet $\Sigma$ besteht, die von einer deterministischen Turingmaschine in polynomialer Zeit akzeptiert werden. Das heißt, zu $A$ gibt es eine deterministische Turingmaschine $M$ und eine $p$-beschränkte Funktion[1] $f\colon \mathbb{N} \to \mathbb{N}$, so daß $M$ angesetzt auf $x \in \Sigma^n$ nach höchstens $f(n)$ Schritten entschieden hat, ob $x$ zu $A$ gehört oder nicht; im ersten Fall gibt $M$ eine 1, ansonsten eine 0 aus. $M$ berechnet also die charakteristische Funktion $\chi_A\colon \Sigma^* \to \{0,1\}$, wobei $\Sigma^* := \cup_{n\geq 0}\Sigma^n$.

Neben einer Vielzahl von effizient lösbaren Problemen traten in der Praxis vermehrt Probleme in den Vordergrund, die sich allen Anstregungen, sie effizient zu lösen, widersetzten. Viele derartige Probleme hatten aber eins gemeinsam: bekam man eine im Vergleich zur Eingabelänge kurze Lösung "verraten", so hatte man es nicht schwer, diese Lösung als solche zu verifizieren. Beispiele derartiger Probleme sind das Erfüllbarkeitsproblem der Aussagenlogik, oder die Frage nach einem Hamiltonkreis in einem Graphen. Dies führte zur Definition der Komplexitätsklasse **NP**.

**Definition 1** *Eine Sprache $A \subseteq \Sigma^*$ gehört zu* **NP** *gdw. es eine $p$-beschränkte Funktion $t\colon \mathbb{N} \to \mathbb{N}$ sowie eine Sprache $B \subseteq (\Sigma \sqcup \{\#\})^*$ aus* **P** *gibt, so daß*

$$\forall n \in \mathbb{N}\ \forall x \in \Sigma^n\ (x \in A \Leftrightarrow \exists e \in \Sigma^{t(n)} :\ x\#e \in B).$$

Hat $x \in A$ die Länge $n$ und ist $e \in \Sigma^{t(n)}$ mit $x\#e \in B$, so ist $e$ ein kurzer Zeuge für die Zugehörigkeit von $x$ zu $A$ und die Möglichkeit, $x\#e \in B$ in polynomialer Zeit zu entscheiden, kann man als effiziente Verifikation von $x \in A$ ansehen. Die obige Äquivalenz kann man auch so umformulieren:

$$\chi_A(x) = \bigvee_{e\in\Sigma^{t(n)}} \chi_B(x\#e).$$

Nun kommen wir zu den algebraischen Komplexitätsklassen und leiten zunächst von den bisherigen Objekten, nämlich den Sprachen, über zu den algebraischen Objekten.

[1]Eine Funktion $f\colon \mathbb{N} \to \mathbb{N}$ heißt *$p$-beschränkt*, wenn sie von einer Polynomfunktion majorisiert wird.

Es sei $A$ eine Sprache über dem Alphabet $\Sigma = \{0, 1\}$. Dieses $A$ kann man ansehen als Folge von Indikatorfunktionen $f_n$, wobei $f_n \colon \{0,1\}^n \to \{0,1\}$ der Indikator von $A \cap \Sigma^n$ in $\Sigma^n$ ist. Zu dieser Funktionenfolge gehört die Partition $A = \sqcup_{n \geq 0}(A \cap \Sigma^n)$. Es wird für unsere Zwecke bequem sein, mit dem Parameter $n$ etwas großzügiger umzugehen. Denken wir etwa an die Sprache $A$ aller invertierbaren Matrizen über dem Körper $\mathbb{F}_2$ aus zwei Elementen, so ist diese in natürlicher Weise partitioniert als $A = \sqcup A_n$, wobei $A_n = \mathrm{GL}(n, 2)$ ist. Demnach ist hier $f_n \colon \{0,1\}^{n^2} \to \{0,1\}$. Da aber die Klasse der $p$-beschränkten Funktionen unter Komposition abgeschlossen ist, können wir uns diese Freiheit erlauben. Schließlich beachte man, daß jede $n$-stellige Boolesche Funktion sich repräsentieren läßt durch ein Polynom $F_n \in \mathbb{F}_2[X_1, \ldots, X_n]$ von höchstens linearem Grad in jeder Variablen, insbesondere ist der Grad von $F_n$ polynomial beschränkt in $n$, sogar $\deg F_n \leq n$.

Bevor wir zur Definition der Komplexitätsklasse **VP** kommen, legen wir noch zwei Bezeichnungsweisen fest. Für ein multivariates Polynom $f$ über dem Körper $k$ sei $v(f)$ die Minimalzahl von Unbestimmten, von denen $f$ abhängt, o.B.d.A. sei $f \in k[X_1, \ldots, X_{v(f)}]$. (Manchmal werden wir auch andere Mengen zum Indizieren der Unbestimmten nehmen.) Im folgenden arbeiten wir der Einfachheit halber mit straight-line Programmen, bei denen nur addiert, multipliziert und skalarmultipliziert werden darf; die zum Kostenmaß $c$ mit $c(\omega) = 1$ für $\omega \in k \cup \{+, *\}$ gehörige Komplexität eines multivariaten Polynoms $f$ (modulo $I = k \cup \{X_1, X_2, \ldots\}$) bezeichnen wir mit $L(f)$.

**Definition 2** *Es sei $k$ ein Körper, $X_1, X_2, \ldots$ Unbestimmte über $k$.*

- *Eine Folge $f = (f_n)_{n \geq 1}$ multivariater Polynome über $k$ heißt eine* $p$-Familie *gdw. die Funktionen $n \mapsto v(f_n)$ und $n \mapsto \deg f_n$ beide $p$-beschränkt sind.*
- *Eine $p$-Familie $f = (f_n)$ heißt* $p$-berechenbar, *gdw. $n \mapsto L(f_n)$ $p$-beschränkt ist.*
- **VP** $=$ **VP**(nonuniform; $k$) *bezeichnet Valiants Klasse aller $p$-berechenbaren Familien über $k$.*

Wir machen einige Anmerkungen zur Definition. Die Einschränkung auf *p-Familien* ist ein Gebot der Fairness. Dadurch wird der unerwünschte Effekt ausgeschaltet, daß durch zu schnell wachsenden Grad oder durch zu schnell wachsende Unbestimmtenanzahl gewisse Polynomfamilien "komplexer" erscheinen als sie in Wirklichkeit sind. *Nichtuniform* bedeutet hier etwa: die Mitglieder der Familie $(f_n)$ müssen nicht notwendigerweise einheitlich durch eine Turingmaschine beschreibbar sein. Wir bemerken noch, daß wir dieselbe Komplexitätsklasse **VP** erhalten, wenn auch Divisionen zugelassen sind und alle Operationen gezählt werden. (Dies folgt aufgrund eines Satzes von Strassen [13].)

**Beispiele.** Folgende $p$-Familien liegen in **VP**:

- $SUM := (SUM_n)_{n\geq 1}$, wobei $SUM_n := X_1 + \ldots + X_n$.
- $PROD := (PROD_n)_{n\geq 1}$, wobei $PROD_n := X_1 \cdots X_n$.
- $POWERSUM := (POWERSUM_n)$, wobei $POWERSUM_n := \sum_{i=1}^n X_i^n$.
- $DET := (DET_n)$; dies folgt mittels Gaußelimination sowie der oben erwähnten Tatsache, daß Divisonen zugelassen werden können.

Wenn wir in $\chi_A(x) = \vee_{e\in\{0,1\}^{t(n)}} \chi_B(x\#e)$ die Disjunktion über $e$ durch eine Summe über $e$ ersetzen, kommen wir zu folgendem Analogon von **NP**.

**Definition 3** • *Eine p-Familie $f = (f_n)$ von Polynomen über $k$ heißt p-definierbar, gdw. es eine p-berechenbare Familie $g = (g_n) \in \mathbf{VP}$ gibt, so daß stets $t(n) := v(g_n) - v(f_n) \geq 0$ ist und $f_n(X) = \sum_{e\in\{0,1\}^{t(n)}} g_n(X, e)$ gilt. (Wir treffen die Konvention $X := (X_1, \ldots, X_{v(f_n)})$ und setzen $f_n(X) = g_n(X)$, falls $t(n) = 0$.)*

- $\mathbf{VNP} = \mathbf{VNP}(\text{nonuniform}; k)$ *bezeichnet Valiants Klasse aller p-definierbaren Familien über $k$.*

Offenbar ist $\mathbf{P} \subseteq \mathbf{NP}$ und $\mathbf{VP} \subseteq \mathbf{VNP}$. Wir geben im folgenden eine $p$-Familie in **VNP** an, von der vermutet wird, daß sie nicht in **VP** liegt, wenn $\operatorname{char} k \neq 2$. (Im Fall der Charakteristik 2 ist $DET = PER$, also $PER \in \mathbf{VP}$.)

**Proposition 4** *$PER \in \mathbf{VNP}$.*

Beweis. $PER$ ist eine $p$-Familie, denn die Variablenanzahlfunktion $n \mapsto n^2$ und die Gradfunktion $n \mapsto n$ sind $p$-beschränkt. Als nächstes geben wir eine Familie $g = (g_n)$ in **VP** an, wobei $g_n = g_n(X, Y)$ ein Polynom in $2n^2$ Unbestimmten $X_{ij}$ und $Y_{ij}$ über $k$ ist mit $PER_n = \sum_{e\in\{0,1\}^{n\times n}} g_n(X, e)$. Die Polynome $g_n$ sind definiert durch

$$g_n(X,Y) := \underbrace{\underbrace{\left(\prod_{i=\ell \,\Leftrightarrow\, j\neq m} (1 - Y_{ij}Y_{\ell m})\right)}_{=:\alpha_n(Y)} \cdot \underbrace{\left(\prod_{i=1}^n \sum_{j=1}^n Y_{ij}\right)}_{=:\beta_n(Y)}}_{=:\gamma_n(Y)} \cdot \underbrace{\left(\prod_{i=1}^n \sum_{j=1}^n X_{ij}Y_{ij}\right)}_{=:\mu_n(X,Y)}.$$

Dann ist $g = (g_n) \in \mathbf{VP}$, denn $v(g_n) = 2n^2$, $\deg g_n = O(n^3)$ und die Komplexität von $g_n$ ist $O(n^3)$. Weiter zeigt man leicht für alle $e \in \{0,1\}^{n\times n}$:

- $\alpha_n(e) \neq 0$ gdw. jede Zeile und jede Spalte von $e$ höchstens eine Eins enthält.

- Sei $\alpha_n(e) \neq 0$. Dann ist $\beta_n(e) \neq 0$ gdw. jede Zeile von $e$ mindestens eine Eins enthält.
- $\gamma_n(e) \neq 0$ gdw. $e$ eine Permutationsmatrix ist.
- $\gamma_n(e) \in \{0,1\}$.
- $\gamma_n(e) \neq 0$ impliziert $\mu_n(X,e) = \prod_{i=1}^n X_{i\sigma(i)}$, wobei $\sigma$ die zu $e$ gehörige Permutation bezeichnet.
- $PER_n = \sum_{e\in\{0,1\}^{n\times n}} g_n(X,e)$.

Dies beweist $PER \in \mathbf{VNP}$.

In der strukturellen Komplexitätstheorie sieht das weitere Vorgehen typischerweise so aus: Mit geeigneten Reduktionsbegriffen partitioniert man die Komplexitätsklassen in Teilklassen von ungefähr gleich schwierigen Problemen. Nachdem das geschehen ist, ist man insbesondere interessiert an härtesten Problemen innerhalb der großen Komplexitätsklassen. Dies sind die sogenannten vollständigen Probleme. Die nachfolgende Definition präzisiert dies auf eine mögliche Art. (Es gibt auch andere Reduktionsbegriffe.)

**Definition 5** *Es seien $A_1 \subseteq \Sigma_1^*$ und $A_2 \subseteq \Sigma_2^*$ Sprachen.*

- *$A_1$ heißt $p$-*reduzierbar *auf $A_2$ (kurz: $A_1 \leq_p A_2$) gdw. eine $p$-berechenbare Funktion $f\colon \Sigma_1^* \to \Sigma_2^*$ existiert, so daß für alle $x \in \Sigma_1^*$ gilt: $x \in A_1 \Leftrightarrow f(x) \in A_2$.*
- *$A_1$ und $A_2$ heißen $p$-*äquivalent *gdw. $A_1 \leq_p A_2$ und $A_2 \leq_p A_1$.*
- *$A \subseteq \Sigma^*$ heißt* $\mathbf{NP}$-vollständig *gdw. $A \in \mathbf{NP}$ und $B \leq_p A$ gilt, für alle $B \in \mathbf{NP}$.*

Die $\mathbf{NP}$-vollständigen Probleme sind untereinander $p$-äquivalent und bilden gerade die härtesten "Brocken" in $\mathbf{NP}$. Weiterhin gilt für ein beliebiges $\mathbf{NP}$-vollständiges Problem $A$:

$$\mathbf{P} = \mathbf{NP} \Leftrightarrow A \in \mathbf{NP}.$$

Cook [5] war der erste, der von einem natürlichen Problem nachweisen konnte, daß es $\mathbf{NP}$-vollständig ist:

**Satz 6 (Cook)** *Das Erfüllbarkeitsproblem der Aussagenlogik ist $\mathbf{NP}$-vollständig.*

Mittlerweile kennt man hunderte von **NP**-vollständigen Problemen, darunter viele sehr praxisrelevante wie etwa das Problem des Handlungsreisenden, oder das Problem der ganzzahligen linearen Optimierung, siehe z.B. [7, 12].

Jetzt kommen wir zu den entsprechenden Begriffen im algebraischen Kontext. (Beim Reduktionsbegriff sind wir in gewisser Weise restriktiver, was letztendlich aber zu stärkeren Aussagen führt.)

**Definition 7**

- *$f \in k[X_1, \ldots, X_n]$ ist eine* Projektion *von $g \in k[X_1, \ldots, X_m]$ gdw. $f = g(a_1, \ldots, a_m)$ für geeignete $a_1, \ldots, a_m$ aus $k \cup \{X_1, \ldots, X_n\}$.*
- *$f = (f_n)$ heißt eine* $p$-Projektion *von $g = (g_n)$ (kurz: $f \preceq_p g$) gdw. eine $p$-beschränkte Funktion $t$ existiert, so daß $f_n$ eine Projektion von $g_{t(n)}$ ist für alle $n$.*
- *Eine $p$-Familie $g$ über $k$ ist* **VNP**-vollständig *gdw. $g$ in* **VNP** *liegt und jedes $f$ aus* **VNP** *eine $p$-Projektion von $g$ ist.*

Offenbar sind **VP** und **VNP** abgeschlossen unter $p$-Projektionen. Der folgende Satz zeigt, daß die Permanentenfamilie zu den schwierigsten Familien in **VNP** gehört.

**Satz 8 (Valiant)** *PER ist* **VNP***-vollständig über jedem Körper der Charakteristik ungleich* 2.

Valiant [15] hat weiterhin gezeigt, daß die Familie $HC = (HC_n)$ der Hamiltonzyklus-polynome vollständig über *jedem* Körper ist (siehe auch [8]). Um das Besondere an der Vollständigkeit der Permanentenfamilie herauszustellen, geben wir zunächst eine etwas andere Charakterisierung der Komplexitätsklasse **NP**. Es sei $t\colon \mathbb{N} \to \mathbb{N}$ eine $p$-beschränkte Funktion und $R \subset \Sigma^* \times \Sigma^*$ eine Relation mit der Eigenschaft, daß für $(x, y) \in \Sigma^n \times \Sigma^m$ aus $R(x, y)$ stets $m \leq t(n)$ folgt. Weiterhin sei $\{x\#y \mid R(x, y)\} \in \mathbf{P}$. Dann nennt man $\{x \mid \exists y : R(x, y)\}$ ein ($p$-beschränktes) *Suchproblem* und die Funktion, die jedem $x$ die Binärkodierung der Anzahl aller $y$ mit $R(x, y)$ zuordnet, das zugehörige *Zählproblem*. Die Klasse **NP** besteht nun gerade aus allen Suchproblemen. Jedes $L \in \mathbf{NP}$ liefert ein Zählproblem $\#L$ und $\#\mathbf{P}$ (lies: number **P**) bezeichnet die Klasse aller Zählprobleme, die von zählenden Turingmaschinen in Polynomialzeit berechnet werden können. Auch $\#\mathbf{P}$ enthält vollständige Probleme bezüglich eines geeigneten Reduktionsbegriffs. Es konnte gezeigt werden, daß für viele **NP**-vollständige Probleme $L$ das zugehörige Zählproblem $\#L$ seinerseits vollständig in $\#\mathbf{P}$ ist, siehe z.B. Valiant [14] und Johnson [12]. Valiant [16] machte darüber hinaus die erstaunliche Entdeckung, daß es sogar Probleme *in* **P** gibt, deren zugehörige Zählprobleme $\#\mathbf{P}$-vollständig sind. Das eindrucksvollste Beispiel ist das Problem des perfekten Matchings in bipartiten Graphen, das nach M. Hall [10] in **P** liegt. Das zugehörige Zählproblem, das äquivalent zur Permanentenberechnung von 0-1 Matrizen ist, stellte sich als #**P**-vollständig heraus. Vor diesem Hintergrund sollte der obige Satz von Valiant gesehen werden.

Die Sätze von Cook und Valiant gaben Anlaß zur

**Hypothese von Cook:** $\mathbf{P} \neq \mathbf{NP}$.
**Hypothese von Valiant:** $\mathbf{VP} \neq \mathbf{VNP}$.

Wir werden im letzten Abschnitt eine verschärfte Version der Valiantschen Hypothese auf algebraisch-kombinatorische Weise formulieren, wodurch schnell klar werden wird, wie weit man noch von einem Beweis dieser Hypothese entfernt ist.

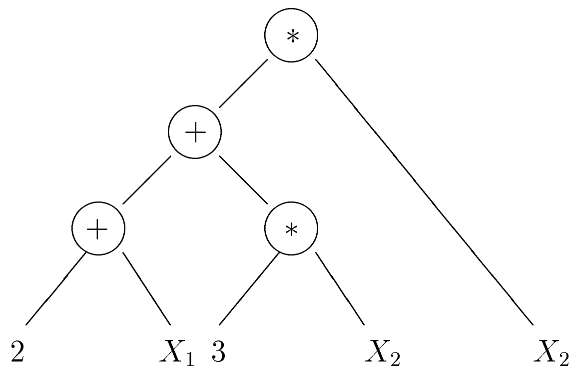
In den restlichen Abschnitten wird eine grobe Beweisskizze des Satzes von Valiant gegeben. Für Einzelheiten verweisen wir auf Kapitel 21 in [3].

# 2 p-Definierbarkeit und Formelgröße

Im ersten Beweisschritt wird eine alternative Charakterisierung der Valiantschen Komplexitätsklasse **VNP** mit Hilfe der Formelgröße gegeben.

Die Menge der arithmetischen Formeln (Ausdrücke) über $I := k \cup \{X_1, \ldots, X_n\}$ ist induktiv wie folgt definiert: jedes Element in $I$ ist eine Formel; sind $\varphi_1$ und $\varphi_2$ Formeln, so auch $(\varphi_1 \circ \varphi_2)$, für $\circ \in \{+, *\}$. Die Größe $E(\varphi)$ einer Formel $\varphi$ ist die Anzahl der $+$ und $*$, die zu ihrem Aufbau benutzt wurden. Jede Formel $\varphi$ stellt in naheliegenderweise ein eindeutig bestimmtes Polynom $\mathrm{val}(\varphi) \in k[X_1, \ldots, X_n]$ dar. Die Formelgröße $E(f)$ von $f \in k[X_1, \ldots, X_n]$ ist die kleinste Größe einer Formel $\varphi$ mit $\mathrm{val}(\varphi) = f$.

Jede Formel $\varphi$ kann man durch einen Baum $T_\varphi$ veranschaulichen. So stellt



z.B. die Formel $\varphi = (((2 + X_1) + (3 * X_2)) * X_2)$ dar. Man beachte, daß eine Formel für das Polynom $f$ als spezielles straight-line Programm angesehen werden kann, bei dem Zwischenresultate nur einmal wiederverwendet werden dürfen; insbesondere gilt

$$L(f) \leq E(f).$$

Jedoch können $L(f)$ und $E(f)$ stark voneinander abweichen, was sich schon aus folgender Bemerkung ergibt:

$$\forall\ f \in k[X_1, \ldots, X_n] \setminus \{0\}:\ \ E(f) \geq \deg(f) - 1.$$

Ist z.B. $f = (f_n)$ mit $f_n := X_1^{2^n}$, so folgt $L(f_n) = n$, aber $E(f_n) = 2^n - 1$. Hier ist $n \mapsto L(f_n)$ $p$-beschränkt, wohingegen $n \mapsto E(f_n)$ exponentiell in $n$ ist. Allerdings ist $f$ keine $p$-Familie, da $n \mapsto \deg f_n$ nicht $p$-beschränkt ist. Offen ist die Frage, ob ein solcher Unterschied innerhalb von **VP** möglich ist.

**Definition 9** • *Eine $p$-Familie $g = (g_n)$ heißt* $p$-ausdrückbar *gdw. $n \mapsto E(g_n)$ $p$-beschränkt ist.*

- $\mathbf{VP}_e = \mathbf{VP}_e(\text{nonuniform}; k)$ *bezeichnet Valiants Klasse aller $p$-ausdrückbaren Familien über $k$.*
- $\mathbf{VNP}_e = \mathbf{VNP}_e(\text{nonuniform}; k)$ *bezeichnet Valiants Klasse aller Familien $f = (f_n)$ über $k$ so daß eine $p$-ausdrückbare Familie $g$ existiert mit $t(n) := v(g_n) - v(f_n) \geq 0$ und $f_n(X) = \sum_{e \in \{0,1\}^{t(n)}} g_n(X, e)$.*

Offenbar ist $\mathbf{VP}_e \subseteq \mathbf{VP} \subseteq \mathbf{VNP}$ und $\mathbf{VNP}_e \subseteq \mathbf{VNP}$. Eine weitere fundamentale Vermutung lautet:

$$\mathbf{VP}_e \neq \mathbf{VP}.$$

Überraschenderweise stimmen die zugehörigen "nichtdeterministischen" Klassen überein:

**Satz 10 (Valiant)** $\mathbf{VNP}_e = \mathbf{VNP}$.

Einen Beweis findet man in Abschnitt 21.2 in [3].

# 3 Universalität von Determinante und Permanente

Im zweiten Beweisschritt wird gezeigt, daß jede $p$-ausdrückbare Polynomfamilie eine $p$-Projektion von $DET$ und $PER$ ist. Genauer gilt folgender Satz.

**Satz 11 (Valiant)** *Hat $f \in k[X_1, \ldots, X_n]$ Formelgröße $u$, so ist $f$ sowohl eine Projektion von $DET_{2u+2}$ als auch eine Projektion von $PER_{2u+2}$.*

Beweisskizze. (von zur Gathen) Wir beschränken uns hier auf $DET$; ähnlich geht man bei $PER$ vor. Es bezeichne $\mathcal{E}$ die Menge aller Formeln über $I := k \cup \{X_1, \ldots, X_n\}$. Man definiert entlang des Formelaufbaus eine Abbildung $\mu \colon \mathcal{E} \to \cup_{s \geq 1} I^{s \times s}$ mit folgenden Eigenschaften für alle $\varphi \in \mathcal{E}$:

(A) $\mathrm{val}(\varphi) = \det(\mu(\varphi))$.

(B) Hat $\varphi$ Formelgröße $u$, so ist die Matrix $\mu(\varphi)$ $s$-reihig, $s = 2u + 2$.

(C) Mit $s = 2u + 2$ aus (B) gibt es $A \in I^{(s-1)\times(s-1)}$, $\alpha \in I^{1\times(s-1)}$, $\beta \in I^{(s-1)\times 1}$, so daß $A$ obere Dreiecksmatrix ist mit Einsen auf der Hauptdiagonalen und

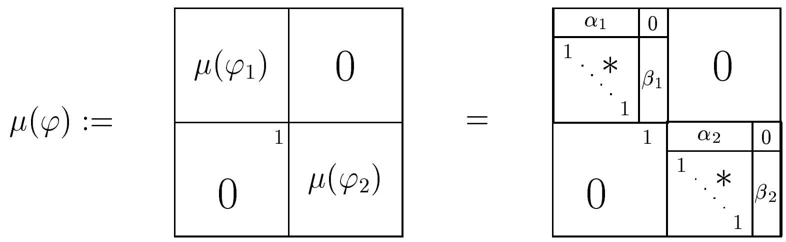
$$\mu(\varphi) = \begin{pmatrix} \alpha & 0 \\ A & \beta \end{pmatrix}.$$

(D) $\mu(\varphi)$ hat in jeder Spalte höchstens einen Eintrag, der nicht in $k$ liegt. Die letzte Spalte enthält keine Unbestimmte.

Konstruktion von $\mu$:

Fall 1. $(u = 0)$ Sei $\varphi \in I$. Dann erfüllt $\mu(\varphi) := \left(\begin{smallmatrix} \varphi & 0 \\ 1 & 1 \end{smallmatrix}\right) \in I^{2\times 2}$ die Bedingungen (A)–(D).

Fall 2. $\varphi = (\varphi_1 * \varphi_2)$. Für $i \in \{1, 2\}$ bezeichne $u_i$ die Formelgröße von $\varphi_i$. Definiere $\mu(\varphi)$ wie folgt:




Dann gelten (C) und (D). Weiterhin ist $u = u_1 + u_2 + 1$ die Größe von $\varphi$ und nach Induktion ist die Größe $s$ von $\mu(\varphi)$ gleich $(2u_1 + 2) + (2u_2 + 2) = 2u + 2$, womit (B) bewiesen ist. Aus der Blockdreiecksgestalt von $\mu(\varphi)$ folgt auch leicht (A).

Fall 3. $\varphi = (\varphi_1 + \varphi_2)$. Wir wenden auf $M_1 := \mu(\varphi_1)$ und $M_2 := \mu(\varphi_2)$ das folgende Lemma an und erhalten $\det(M) = -\det(M_1) - \det(M_2) = -\mathrm{val}(\varphi)$ für die dort konstruierte Matrix $M$. Wir bekommen $\mu(\varphi)$, indem wir zu $M$ eine letzte Zeile und eine vorletzte

Spalte hinzufügen, deren Einträge sämtlich Null sind mit Ausnahme der Kreuzungsstelle, an der eine Eins steht.

**Lemma 12** *Es sei $R$ ein kommutativer Ring. Für $i = 1, 2$ sei $A_i \in R^{d_i \times d_i}$ eine obere Dreiecksmatrix mit Einsen auf der Diagonalen, $\alpha_i \in R^{1 \times d_i}$ und $\beta_i \in R^{d_i \times 1}$. Dann stehen die Determinanten und Permanenten der Matrizen*

$$M_1 := \begin{pmatrix} \alpha_1 & 0 \\ A_1 & \beta_1 \end{pmatrix}, M_2 := \begin{pmatrix} \alpha_2 & 0 \\ A_2 & \beta_2 \end{pmatrix}, M := \begin{pmatrix} \alpha_1 & \alpha_2 & 0 \\ A_1 & 0 & \beta_1 \\ 0 & A_2 & \beta_2 \end{pmatrix}$$

*wie folgt in Beziehung:*

$$\det(M) = (-1)^{d_2} \det(M_1) + (-1)^{d_1} \det(M_2)$$

*und*

$$\mathrm{per}(M) = \mathrm{per}(M_1) + \mathrm{per}(M_2).$$

Der Beweis des Lemmas ergibt sich durch Laplace-Entwicklung nach der $d_1 + 1$-ten Spalte von $M$. Damit ist die Beweisskizze des Universalitätssatzes abgeschlossen. Valiant [15] zeigt mit einer kompakteren Konstruktion, daß im letzten Satz $2u + 2$ sogar durch $u + 3$ ersetzt werden kann.

# 4 Die Vollständigkeit der Permanentenfamilie

Nach den bisherigen Vorbereitungen kommen wir jetzt zur Skizze des eigentlichen Vollständigkeitsbeweises. Wir wissen bereits, daß *PER* in **VNP** liegt. Es bleibt zu zeigen, daß jedes $f \in \mathbf{VNP}$ eine $p$-Projektion von *PER* ist. Wegen $\mathbf{VNP} = \mathbf{VNP}_e$ gibt es zu jedem $f \in \mathbf{VNP}$ ein $g \in \mathbf{VP}_e$ mit $f_n(X) = \sum_e g_n(X, e)$, für alle $n$. Sei $m = v(f_n)$ und $t = v(g_n) - v(f_n)$. Setze $X = (X_1, \ldots, X_m)$ und $Y = (Y_1, \ldots, Y_t) := (X_{m+1}, \ldots, X_{m+t})$. Aufgrund der Universalität der Permanente gibt es eine Matrix $A$ über $k \cup \{X_1, \ldots, X_m\} \cup \{Y_1, \ldots, Y_t\}$ mit $N = 2E(g_n) + 2$ Reihen, für die $g_n(X, Y) = \mathrm{per}(A)$ ist. Weiter können wir nach Eigenschaft (D) im Universalitätsbeweis annehmen, daß $A$ in jeder Spalte höchstens einen Eintrag hat, der eine Unbestimmte ist. Damit ergibt sich die **VNP**-Vollständigkeit von *PER* aus folgendem Resultat.

**Satz 13** *Es sei $k$ ein Körper der Charakteristik $\neq 2$ und $A = A(X, Y)$ eine $N \times N$ Matrix über $k \cup \{X_1, \ldots, X_m, Y_1, \ldots, Y_t\}$, in der pro Spalte höchstens ein Eintrag außerhalb $k$ vorkommt. Dann läßt sich eine quadratische Matrix $A'$ über $k \cup \{X_1, \ldots, X_m\}$ mit $N' \leq 10N$ Zeilen angeben, so daß $\mathrm{per}(A') = \sum_{e \in \{0,1\}^t} \mathrm{per} A(X, e)$.*

Beweisskizze. Die zu konstruierende Matrix $A'$ hat eine Block- und eine Feinstruktur. Blockeinträge ungleich einer Nullmatrix gibt es in $A'$ höchstens in der ersten Blockzeile, in der ersten Blockspalte, sowie auf der Blockdiagonalen. Das folgende Schaubild verdeutlicht die Blockstruktur für den Fall $t = 2$:

$$A' = \begin{pmatrix} A_0 & \mathcal{Y}_{01} & \mathcal{Y}_{02} \\ \mathcal{Y}_{10} & \mathcal{Y}_1 & 0 \\ \mathcal{Y}_{20} & 0 & \mathcal{Y}_2 \end{pmatrix}.$$

Dabei geht $A_0$ aus $A$ hervor, indem man alle $Y_i$ durch 0 ersetzt. Bei der Feinstruktur spielt die Valiant-Matrix

$$V = \begin{pmatrix} 0 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 3 & 0 \end{pmatrix}$$

eine zentrale Rolle. Dies liegt an folgenden Eigenschaften (dabei bezeichne $V[R|C]$ die Matrix $V$ ohne die Zeilen $r \in R$ und Spalten $c \in C$):

$$\mathrm{per}(V) = \mathrm{per}(V[1|1]) = \mathrm{per}(V[4|4]) = \mathrm{per}(V[1,4|1,4]) = 0$$

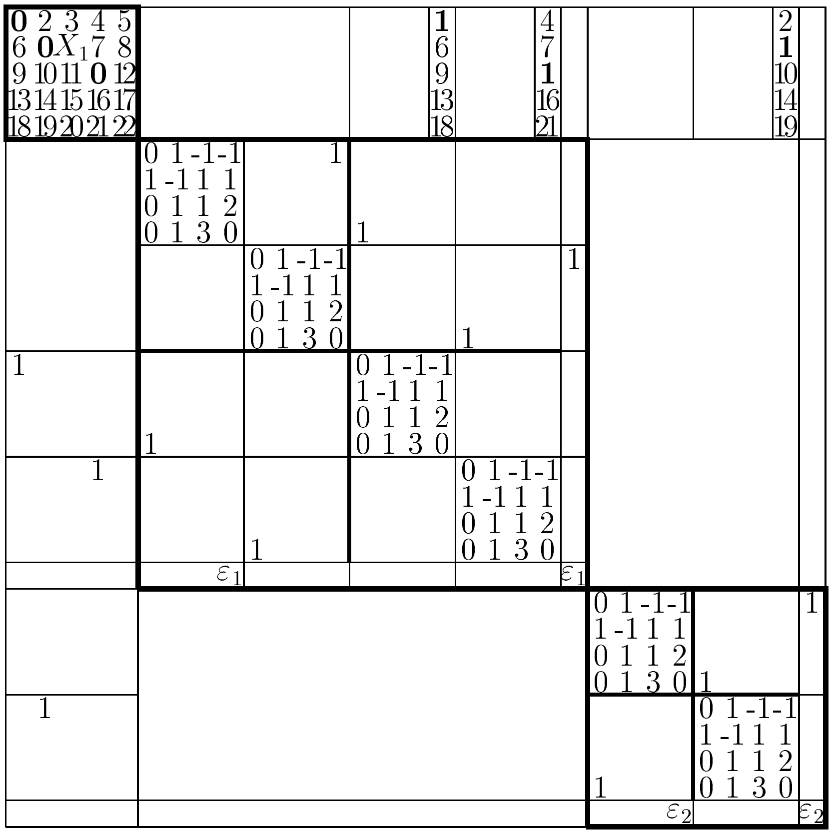
und

$$\mathrm{per}(V[1|4]) = \mathrm{per}(V[4|1]) = 4.$$

Wir erläutern die Feinstruktur von $A'$ anhand des Beispiels:

$$A = A(X_1, Y_1, Y_2) = \begin{pmatrix} Y_1 & 2 & 3 & 4 & 5 \\ 6 & Y_2 & X_1 & 7 & 8 \\ 9 & 10 & 11 & Y_1 & 12 \\ 13 & 14 & 15 & 16 & 17 \\ 18 & 19 & 20 & 21 & 22 \end{pmatrix}.$$

(Hier ist also $m = 1$ und $t = 2$.) Die zugehörige Matrix $A'$ sieht so aus (dabei ist $\varepsilon_1 = 4^{-4}$ und $\varepsilon_2 = 4^{-2}$, was wegen der Voraussetzung $\mathrm{char}\, k \neq 2$ Sinn macht!):



Das Ergebnis $\mathrm{per}(A') = \sum_{e \in \{0,1\}^2} \mathrm{per} A(X_1, e_1, e_2)$ ergibt sich nun mittels einer verallgemeinerten Laplace-Entwicklung unter Verwendung der Blockstruktur der Matrix $A'$. Dabei steuert die Zeile mit den beiden $\varepsilon_1$'s die Summation $e_1 \in \{0, 1\}$: das linke $\varepsilon_1$ liefert den Beitrag zu "$e_1 = 1$" das andere den zu "$e_1 = 0$". Entsprechendes gilt für die beiden $\varepsilon_2$'s.

Das mag an groben Hinweisen genügen. Einzelheiten findet man im Abschnitt 21.4 von [3].

# 5 Die erweiterte Valiantsche Hypothese

Wir wollen in diesem letzten Abschnitt die Valiantsche Hypothese verschärfen. Dazu arbeiten wir mit $p$-Familien, deren Formelgrößen bzw. Komplexitäten quasi-polynomial wachsen dürfen; das ist schneller als polynomial, aber weniger schnell als exponentiell.

**Definition 14** • *Eine Funktion* $t\colon \mathbb{N} \to \mathbb{N}$ *heißt* quasi-polynomial beschränkt *(qp-*beschränkt*), wenn es eine positive Konstante $c$ gibt mit $t(n) \leq n^{O(\log^c n)}$.*

- *Eine p-Familie* $f = (f_n)$ *über* $k$ *heißt qp-*berechenbar *(bzw. qp-ausdrückbar) gdw.* $n \mapsto L(f_n)$ *(bzw.* $n \mapsto E(f_n)$*) qp-beschränkt ist.*

- $\mathbf{VQP} = \mathbf{VQP}(k; \textit{nonuniform})$ *(bzw.* $\mathbf{VQP}_e = \mathbf{VQP}_e(k; \textit{nonuniform})$*) bezeichnet Valiants Klasse aller qp-berechenbaren (bzw. qp-ausdrückbaren) Familien über* $k$.

Offenbar ist $\mathbf{VP} \subseteq \mathbf{VQP}$ und $\mathbf{VP}_e \subseteq \mathbf{VQP}_e$. Die folgende Vermutung verallgemeinert die Hypothese $\mathbf{VNP} \setminus \mathbf{VP} \neq \emptyset$.

**Erweiterte Hypothese von Valiant:** $\mathbf{VNP} \setminus \mathbf{VQP} \neq \emptyset$ über jedem Körper.

Mit den folgenden Ausführungen soll skizziert werden, daß diese Vermutung äquivalent ist zur Aussage: $\mathbf{VNP}_e \setminus \mathbf{VQP}_e \neq \emptyset$ über einem beliebigen Körper. Da wir bereits wissen, daß $\mathbf{VNP}_e = \mathbf{VNP}$ ist, genügt es, folgenden Satz zu zeigen.

**Satz 15** $\mathbf{VQP}_e = \mathbf{VQP}$ *über einem beliebigen Körper.*

Dieser Satz ergibt sich aus Zusammenhängen zwischen den verschiedenen Komplexitätsmaßen Formelgröße $E(f)$, Komplexität $L(f)$ und der *Tiefe* $D(f)$ eines Polynoms $f$. Die Tiefe kann interpretiert werden als die die minimale parallele Berechnungszeit, und im folgenden präzisieren wir kurz diesen Begriff. Jedem straight-line Programm $\Gamma = (\Gamma_1, \ldots, \Gamma_r)$, das Eingaben der Länge $n$ erwartet, kann man einen Digraphen zuordnen, dessen Knotenmenge $\{-n+1, \ldots, r\}$ ist. Eine Anweisung $\Gamma_i = (\omega_i; \alpha, \beta)$ steuert zwei Kanten, nämlich $(\alpha, i)$ und $(\beta, i)$ bei, während die Skalarmultiplikationsanweisung $\Gamma_i = (\omega_i; \alpha)$ nur die Kante $(\alpha, i)$ beiträgt. Die Tiefe $D(\Gamma)$ von $\Gamma$ ist die maximale Länge eines Weges im gerade definierten Digraphen zu $\Gamma$. Die *Tiefe* (depth) des Polynoms $f$ ist definiert durch

$$D(f) := \min\{D(\Gamma) \mid \Gamma \text{ berechnet } f\}.$$

Entsprechend definiert man die Tiefe $T(\varphi)$ einer Formel $\varphi$ und gelangt so zum Begriff der *Formeltiefe* $T(f)$ von $f$:

$$T(f) := \min\{T(\varphi) \mid \mathrm{val}(\varphi) = f\}.$$

Es ist eine empfehlenswerte Übung zu zeigen, daß $D(f) = T(f)$ ist. Während die untere Schranke des folgenden Satzes fast trivial ist (ebenfalls eine sinnvolle Übung!), ergibt sich die obere Schranke durch geschickte Anwendung des goldenen Schnitts; daher tritt dort auch die Zahl $\epsilon := (1 + \sqrt{5})/2$ auf.

**Satz 16 (Brent [2])** *Für ein n-variates Polynom* $f$ *vom Grad* $d \geq 2$ *gilt:*

$$\log(E(f) + 1) \leq D(f) \leq \frac{2}{\log \epsilon} \log(E(f)) + 1.$$

Die Tatsache, daß $D = \Theta(\log E)$ ist, untermauert die Vermutung, daß $\mathbf{VP}_e \neq \mathbf{VP}$ gilt, denn $\mathbf{VP}_e = \mathbf{VP}$ würde implizieren, daß für jedes $f \in \mathbf{VP}$ (insbesondere auch für $f = DET$) eine Konstante $c$ existiert mit $D(f_n) \leq c \log n$, für alle $n$. Das liegt aber jenseits unserer Vorstellungskraft.

In den Beweis von $\mathbf{VQP}_e = \mathbf{VQP}$ geht schließlich noch das folgende fundamentale Resultat der parallelen Komplexitätstheorie ein, das auf Hyafil [11] sowie Valiant, Skyum, Berkowitz und Rackoff [18] zurückgeht.

**Satz 17** *Es gibt es eine universelle Konstante $c$, so daß für jedes $n$-variate Polynom $f$ vom Grad $d \geq 1$ über $k$ gilt:*

$$D(f) \leq c(\log(dL(f)) \log d + \log n).$$

*Darüber hinaus kann $f$ von einem straight-line Programm der Länge $O(d^6 L(f)^3)$ und Tiefe $O(\log(dL(f)) \log d + \log n)$ berechnet werden.*

Aus den letzten beiden Sätzen ergibt sich nach leichter Rechnung die Gleichheit $\mathbf{VQP} = \mathbf{VQP}_e$.

Zum Vergleich von $DET$ und $PER$ diskutieren wir jetzt Vollständigkeitsresultate für $\mathbf{VQP}$ auf der Basis eines etwas großzügigeren Reduktionsbegriffs.

**Definition 18**

- *Seien $f = (f_n)$ und $g = (g_n)$ $p$-Familien über $k$. Dann heißt $f$ eine* $qp$-Projektion *von $g$ gdw. es eine $qp$-beschränkte Funktion $t$ gibt, so daß für jedes $n$ das Polynom $f_n$ eine Projektion von $g_{t(n)}$ ist.*
- *Eine Familie $g$ heißt* $\mathbf{VQP}$-vollständig *gdw. $g \in \mathbf{VQP}$ und jedes $f \in \mathbf{VQP}$ eine $qp$-Projektion von $g$ ist.*

Offenbar ist $\mathbf{VQP}$ abgeschlossen unter $qp$-Projektionen.

**Satz 19** *$DET$ ist $\mathbf{VQP}$-vollständig.*

Beweis. Wir wissen bereits, daß $DET \in \mathbf{VP} \subseteq \mathbf{VQP}$. Nun sei $f \in \mathbf{VQP} = \mathbf{VQP}_e$. Dann ist $n \mapsto E(f_n)$ $qp$-beschränkt und aufgrund der Universalität der Determinante ist $f_n$ eine Projektion von $DET_{2E(f_n)+2}$. Also ist $f$ eine $qp$-Projektion von $DET$.

Nun kommen wir zu einem abschließenden Vergleich von $DET$ und $PER$. Aufgrund von Satz 17 wissen wir, daß $DET_n$ durch ein straight-line Programm polynomialer Länge und Tiefe $O(\log^2 n)$ berechnet werden kann. (Für direkte Konstruktionen verweisen wir auf Csanky [6], Berkowitz [1] und Chistov [4].) Zusammen mit $D = \Theta(\log E)$ (Satz von Brent) ergibt das

$$E(DET_n) = 2^{O(\log^2 n)}.$$

Im Vergleich dazu ergibt sich für die Permanente aus der Formel von Ryser nur die folgende obere Schranke:

$$E(PER_n) = O(n^2 2^n).$$

Wegen Satz 8 und Satz 19 ist die erweiterte Valiantsche Hypothese (in Charakteristik $\neq 2$) äquivalent zur folgenden rein algebraisch-kombinatorischen Aussage.

**Erweiterte Hypothese von Valiant:** $PER$ ist keine $qp$-Projektion von $DET$.

Das bisher beste Resultat, was in diese Richtung geht, besagt folgendes:

**Satz 20 (von zur Gathen [9])** *$PER_n$ ist keine Projektion von $DET_m$, falls $m < \sqrt{2}n$.*

Zum Beweis der erweiterten Valiantschen Hypothese müßte dieses Resultat um astronomische Größenordungen verbessert werden! Vielleicht kann die Kombinatorik hier der Komplexitätstheorie weiterhelfen.

# Literatur

[1] S. Berkowitz. On computing the determinant in small parallel time using a small number of processors. *Inf. Proc. Letters*, 18:147–150, 1984.

[2] R.P. Brent. The complexity of multiprecision arithmetic. In *Proc. Seminar on Compl. of Comp. Problem Solving.* Brisbaine, 126–165, 1975.

[3] P. Bürgisser, M. Clausen, and M.A. Shokrollahi. *Algebraic Complexity Theory, Grundlehren der mathematischen Wissenschaften*, Bd. 315. Springer Verlag, 1996.

[4] A.L. Chistov. Fast parallel calculation of the rank of matrices over a field of arbitrary characteristic. In *Fundamentals of Computation Theory*, number 199 in *Lecture Notes in Computer Science*, Springer-Verlag, 63–69, 1985.

[5] S.A. Cook. The complexity of theorem proving procedures. In *Proc. 3rd ACM STOC*, 151–158, 1971.

[6] L. Csanky. Fast parallel matrix inversion algorithms. *SIAM J. Comp.*, 5:618–623, 1976.

[7] M.R. Garey und D.S. Johnson. *Computers and intracrability: a guide to the theory of* **NP**-*completeness.* W.H. Freeman and Company, New York, 1979.

[8] J. von zur Gathen. Feasible arithmetic computations: Valiant's hypothesis. *J. Symb. Comput.*, 4:137–172, 1987.

[9] J. von zur Gathen. Permanent and determinant. *Lin. Alg. Appl.*, 96:87–100, 1987.

[10] M. Hall. An algorithm for distinct representatives. *Amer. Math. Monthly*, 63:716–717, 1956.

[11] L. Hyafil. On the parallel evaluation of multivariate polynomials. *SIAM J. Comp.*, 8:120–123, 1979.

[12] D.S. Johnson. A catalog of complexity classes. In *J. van Leeuwen (ed.): Handbook of Theoretical Computer Science, volume A, chapter 2, 61–161.* Elsevier Science Publishers B. V., 1990.

[13] V. Strassen. Vermeidung von Divisionen. *Crelles J. Reine Angew. Math.*, 264:184–202, 1973.

[14] L.G. Valiant. The complexity of enumeration and reliability problems. *SIAM J. Comp.*, 8:410–421, 1979.

[15] L.G. Valiant. Completeness classes in algebra. In *Proc. 11th ACM STOC*, 249–261, 1979.

[16] L.G. Valiant. The complexity of computing the permanent. *Theoret. Comput. Sc.*, 8:189–201, 1979.

[17] L.G. Valiant. Reducibility by algebraic projections. In *Logic and Algorithmic: an International Symposium held in Honor of Ernst Specker*, volume 30, pp. 365–380, 1982.

[18] L.G. Valiant, S. Skyum, S. Berkowitz und C. Rackoff. Fast parallel computation of polynomials using few processors. *SIAM J. Comp.*, 12(4):641–644, 1983.