# Approximate Interpolation

## on Finite Sets of Points

Martin Kreuzer

Fakultät für Informatik und Mathematik

Universität Passau

martin.kreuzer@uni-passau.de

CoCoA School Osnabrück

June 14, 2013

# Contents

# 1 – The Approximate Interpolation Problem

**I am approximately 96.694444% in love with you.**
**Of course, that's just a rough estimate.**
(Jarod Kintz)

**Question 1:** Given a finite set of points $\mathbb{X} = \{p_1, \ldots, p_s\} \subset \mathbb{R}^n$ and numbers $a_1, \ldots, a_s \in \mathbb{R}$, how can we find a **simple** polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ such that

$$f(p_i) \approx a_i \qquad \text{for} \quad i = 1, \ldots, s\,?$$

**Question 2:** What is the precise meaning of "simple" and of $f(p_i) \approx a_i$ here?

**Question 3:** The polynomial $f$, if it exists, is not uniquely determined. If $g \in \mathbb{R}[x_1, \ldots, x_n]$ is a polynomial such that $g(p_i) \approx 0$ for all $i$, is $f + g$ also a solution of **Problem 1**?

**Question 4:** When does a polynomial **vanish approximately** at $\mathbb{X}$?

**Preliminary Definition.** Let $\varepsilon > 0$ be a given **threshold number**. We say that $f \in P = \mathbb{R}[x_1, \ldots, x_n]$ **vanishes $\varepsilon$-approximately** at $\mathbb{X}$ if $|f(p_i)| < \varepsilon$ for $i = 1, \ldots, s$.

**Question 5:** What is the structure of the set of all polynomials which vanish $\varepsilon$-approximately at $\mathbb{X}$?

**Question 6:** Don't all polynomials with very small coefficients vanish $\varepsilon$-approximately at $\mathbb{X}$?

Clearly, we need to measure the **size** of a polynomial. In other words, we need a topology on $\mathbb{R}[x_1, \ldots, x_n]$.

**Definition 1.1** Let $f = c_1 t_1 + \cdots + c_s t_s \in P$, where $c_1, \ldots, c_s \in \mathbb{R} \setminus \{0\}$ and $t_1, \ldots, t_s \in \mathbb{T}^n$. Then the number $\|f\| = \|(c_1, \ldots, c_s)\|$ is called the **(Euclidean) norm** of $f$.

It is easy to check that this definition turns $P$ into a normed vector space. Now it is reasonable to consider the following condition.

**Definition 1.2** Let $\varepsilon > 0$ be a given **threshold number**. We say that $f \in P = \mathbb{R}[x_1, \ldots, x_n]$ **vanishes $\varepsilon$-approximately** at $\mathbb{X}$ if $|f(p_i)| < \varepsilon \cdot \|f\|$ for $i = 1, \ldots, s$.

**Preliminary Definition.** An ideal $I \subseteq P$ is called an **$\varepsilon$-approximate vanishing ideal** of $\mathbb{X}$ if there exists a system of generators $\{f_1, \ldots, f_r\}$ of $I$ such that every $f_i$ vanishes $\varepsilon$-approximately at $\mathbb{X}$ for $i = 1, \ldots, r$.

**Question 7:** Are there non-trivial $\varepsilon$-approximate vanishing ideals?

A set of $\varepsilon$-approximately vanishing polynomials will almost always generate the **unit ideal**.

**Question 8:** Are there $\varepsilon$-approximately vanishing polynomials such that **close-by** there exist polynomials defining a set of points?

A positive answer will be given by **approximate border bases**.

## 2 – Characterizing Exact Border Bases

**When they started to prove even the simplest claims, many turned out to be wrong**

(Bertrand Russell)

$P = K[x_1, \ldots, x_n]$ polynomial ring over a field $K$

$I \subseteq P$ zero-dimensional polynomial ideal (i.e. $\dim_K(P/I) < \infty$)

$\mathbb{T}^n = \{x_1^{\alpha_1} \cdots x_n^{\alpha_n} \mid \alpha_i \geq 0\}$ monoid of terms

## Definition of Border Bases

**Definition 2.1** **(a)** A (finite) set $\mathcal{O} \subset \mathbb{T}^n$ is called an **order ideal** if $t \in \mathcal{O}$ and $t' \mid t$ implies $t' \in \mathcal{O}$.

**(b)** Let $\mathcal{O}$ be an order ideal. The set $\partial\mathcal{O} = (x_1\mathcal{O} \cup \cdots \cup x_n\mathcal{O}) \setminus \mathcal{O}$ is called the **border** of $\mathcal{O}$.

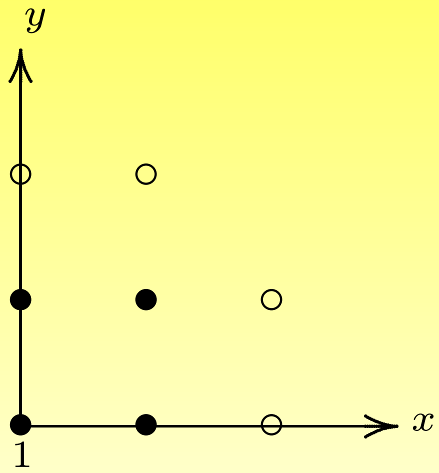
**(c)** Let $\mathcal{O} = \{t_1, \ldots, t_\mu\}$ be an order ideal and $\partial\mathcal{O} = \{b_1, \ldots, b_\nu\}$ its border. A set of polynomials $\{g_1, \ldots, g_\nu\} \subset I$ of the form

$$g_j = b_j - \sum_{i=1}^{\mu} c_{ij} t_i$$

with $c_{ij} \in \mathbb{R}$ and $t_i \in \mathcal{O}$ is called an **$\mathcal{O}$-border prebasis** of $I$.

**(d)** An $\mathcal{O}$-border prebasis of $I$ is called an **$\mathcal{O}$-border basis** of $I$ if the residue classes of the terms in $\mathcal{O}$ are a $K$-vector space basis of $P/I$.

# A Picture of an Order Ideal and its Border



## Neighbors

**Definition 2.2** Let $b_i, b_j \in \partial\mathcal{O}$ be two distinct border terms.

**(a)** The border terms $b_i$ and $b_j$ are called **next-door neighbors** if $b_i = x_k\, b_j$ for some $k \in \{1, \ldots, n\}$.

**(b)** The border terms $b_i$ and $b_j$ are called **across-the-street neighbors** if $x_k\, b_i = x_\ell\, b_j$ for some $k, \ell \in \{1, \ldots, n\}$.

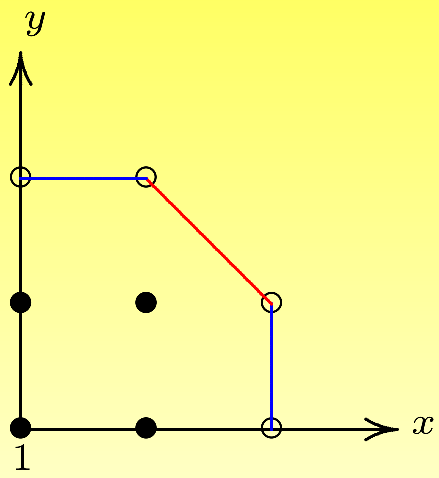
**(c)** The border terms $b_i$ and $b_j$ are called **neighbors** if they are next-door neighbors or across-the-street neighbors.

**(d)** The graph whose vertices are the border terms and whose edges are given by the neighbor relation is called the **border web** of $\mathcal{O}$.

**Example 2.3** The border of $\mathcal{O} = \{1,\, x,\, y,\, xy\}$ is
$\partial\mathcal{O} = \{x^2,\, x^2y,\, xy^2,\, y^2\}$. Here the border web looks as follows:

$(x^2,\, x^2y)$ and $(y^2, xy^2)$ are next-door neighbor pairs

$(x^2y, xy^2)$ is an across-the-street neighbor pair.




**Proposition 2.4** *The border web is connected.*

## The Buchberger Criterion for Border Bases

A border prebasis $G = \{g_1, \ldots, g_\nu\}$ is a border basis if and only if for every neighboring pair $(b_i, b_j)$ of border terms such that $b_i = x_k\, b_j$ resp. $x_k\, b_i = x_\ell\, b_j$ the corresponding **S-polynomial**

$$S_{ij} \;=\; g_i - x_k\, g_j \qquad \text{resp.} \qquad S_{ij} \;=\; x_k\, g_i - x_\ell\, g_j$$

satisfies $\mathrm{NR}_G(S_{ij}) = 0$.

## Border Bases and Multiplication Matrices

For $r \in \{1, \ldots, n\}$, we define the $r$-th **formal multiplication matrix** $A_r$ as follows:

Multiply $t_i \in \mathcal{O}$ by $x_r$. If $x_r t_i = b_j$ is in the border of $\mathcal{O}$, rewrite it using the prebasis polynomial $g_j = b_j - \sum_{k=1}^{\mu} c_{kj} t_k$ and put $(c_1, \ldots, c_\mu)$ into the $i$-th column of $A_r$. But if $x_r t_i = t_j$ then put the $j$-th unit vector into the $i$-th column of $A_r$.

**Theorem 2.5** *The set $G$ is the $\mathcal{O}$-border basis of $I$ if and only if the formal multiplication matrices commute, i.e. iff*

$$A_i \, A_j \;=\; A_j \, A_i \qquad \textit{for } 1 \le i < j \le n.$$

## 3 – Approximate Border Bases

**Need to approximate a border?**

**Try Approximate Border Bases !**

(Anonymous)

$\mathcal{O} = \{t_1, \dots, t_\mu\}$ order ideal in $\mathbb{T}^n$

$\partial\mathcal{O} = \{b_1, \dots, b_\nu\}$ border of $\mathcal{O}$

Recall that an $\mathcal{O}$-border prebasis is a set of polynomials $G = \{g_1, \dots, g_\nu\}$ such that

$$g_j \;=\; b_j \;-\; \sum_{i=1}^{\mu} c_{ij}\, t_i$$

where $c_{ij} \in \mathbb{R}$

## Definition of Approximate Border Bases

**Definition 3.1** Let $\varepsilon > 0$.

An $\mathcal{O}$-border prebasis $G = \{g_1, \ldots, g_\nu\}$ is called an **$\varepsilon$-approximate $\mathcal{O}$-border basis** if the following condition is satisfied:

For every pair $(i,j)$ such that $(b_i, b_j)$ are neighbors, we have $\|\operatorname{NR}_{\mathcal{O},G}(S_{ij})\| < \varepsilon$.

**Remark 3.2** If $G = \{g_1, \ldots, g_\nu\}$ is an $\varepsilon$-approximate border basis then the point $(c_{11}, \ldots, c_{\mu\nu})$ in $\mathbb{R}^{\mu\nu}$ given by its coefficients is **close to** the **border basis scheme**.

**Example 3.3** Let $\mathcal{O} = \{1,\, x,\, y,\, xy\}$. Then the set

$$g_1 = x^2 + 0.02xy - 0.01y - 1.01 \qquad g_2 = x^2y + 0.03x - 0.98\,y$$
$$g_3 = xy^2 - 1.02\,x \qquad g_4 = y^2 - 0.99$$

is a 0.1-approximate $\mathcal{O}$-border basis.

For instance, we have $S_{23} = yg_2 - xg_3 = 0.03xy - 0.98y^2 + 1.02x^2$ and $\mathrm{NR}_{\mathcal{O},G}(S_{23}) \approx 0.01xy + 0.01y + 0.05$.

The ideal $I = \langle g_1, g_2, g_3, g_4\rangle$ is the unit ideal, since $g_3 - x\,g_4 = 0.03x$ shows $-g_1 \equiv 0.01y + 1.01 \pmod I$ and $g_4 \equiv 101^2 - 0.99 \pmod I$.

## 4 – Computing Approximate Border Bases

**We all know Linux is great...**
**it does infinite loops in 5 seconds.**
(Linus Torvalds)

**Goal:** Given a set of (approximate) points $\mathbb{X} = \{p_1, \ldots, p_s\}$ in $\mathbb{R}^n$ and $\varepsilon > 0$, find an order ideal $\mathcal{O}$ and an approximate $\mathcal{O}$-border basis $G$ such that the polynomials in $G$ vanish $\varepsilon$-approximately at the points of $\mathbb{X}$.

Notice that, in general,

- we have $\#\mathcal{O} << \#\mathbb{X}$,
- the ideal $\langle G \rangle$ is the unit ideal.

**Theorem 4.1 (The Singular Value Decomposition)**
*Let* $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{R})$*.*

1. *There are orthogonal matrices* $\mathcal{U} \in \mathrm{Mat}_{m,m}(\mathbb{R})$ *and* $\mathcal{V} \in \mathrm{Mat}_{n,n}(\mathbb{R})$ *and a matrix* $\mathcal{S} \in \mathrm{Mat}_{m,n}(\mathbb{R})$ *of the form* $\mathcal{S} = \begin{pmatrix} \mathcal{D} & 0 \\ 0 & 0 \end{pmatrix}$ *such that*

$$\mathcal{A} = \mathcal{U} \cdot \mathcal{S} \cdot \mathcal{V}^{\mathrm{tr}} = \mathcal{U} \cdot \begin{pmatrix} \mathcal{D} & 0 \\ 0 & 0 \end{pmatrix} \cdot \mathcal{V}^{\mathrm{tr}}$$

   *where* $\mathcal{D} = \mathrm{diag}(s_1, \ldots, s_r)$ *is a diagonal matrix.*

2. *In this decomposition, it is possible to achieve* $s_1 \geq s_2 \geq \cdots \geq s_r > 0$*. The numbers* $s_1, \ldots, s_r$ *depend only on* $\mathcal{A}$ *and are called the* **singular values** *of* $\mathcal{A}$*.*

3. *The number $r$ is the* **rank** *of $\mathcal{A}$.*

4. *The matrices $\mathcal{U}$ and $\mathcal{V}$ have the following interpretation:*

$$
\begin{array}{rcl}
\textit{first } r \textit{ columns of } \mathcal{U} & \equiv & \textit{ONB of the column space of } \mathcal{A} \\
\textit{last } m-r \textit{ columns of } \mathcal{U} & \equiv & \textit{ONB of the kernel of } \mathcal{A}^{\mathrm{tr}} \\
\textit{first } r \textit{ columns of } \mathcal{V} & \equiv & \textit{ONB of the row space of } \mathcal{A} \\
 & \equiv & \textit{ONB of the column space of } \mathcal{A}^{\mathrm{tr}} \\
\textit{last } n-r \textit{ columns of } \mathcal{V} & \equiv & \textit{ONB of the kernel of } \mathcal{A}
\end{array}
$$

**Corollary 4.2** *Let* $\mathcal{A} \in \mathrm{Mat}_{m,n}(\mathbb{R})$*, and let* $\varepsilon > 0$ *be given. Let* $k \in \{1, \dots, r\}$ *be chosen such that* $s_k > \varepsilon \geq s_{k+1}$*. Form the matrix* $\widetilde{\mathcal{A}} = \mathcal{U}\,\widetilde{\mathcal{S}}\,\mathcal{V}^{\mathrm{tr}}$ *by setting* $s_{k+1} = \cdots = s_r = 0$ *in* $\mathcal{S}$*. The matrix* $\widetilde{\mathcal{A}}$ *is called the* **singular value truncation** *of* $\mathcal{A}$*.*

1. *We have* $\min\{\|\mathcal{A} - \mathcal{B}\| : \mathrm{rank}(\mathcal{B}) \leq k\} = \|\mathcal{A} - \widetilde{\mathcal{A}}\| = s_{k+1}$*. (Here* $\|\cdots\|$ *denotes the 2-operator norm of a matrix.)*
2. *The vector subspace* $\mathrm{apker}(\mathcal{A}, \varepsilon) = \ker(\widetilde{\mathcal{A}})$ *is the largest dimensional kernel of a matrix whose Euclidean distance from* $\mathcal{A}$ *is at most* $\varepsilon$*. It is called the* **$\varepsilon$-approximate kernel** *of* $\mathcal{A}$*.*
3. *The last* $n - k$ *columns* $v_{k+1}, \dots, v_n$ *of* $\mathcal{V}$ *are an ONB of* $\mathrm{apker}(\mathcal{A}, \varepsilon)$*. They satisfy* $\|\mathcal{A} v_i\| < \varepsilon$*.*

**Theorem 4.3** **(The AVI Algorithm)**
*The following algorithm computes an approximate border basis of an approximate vanishing ideal of a finite set of points* $\mathbb{X} \subseteq \mathbf{[-1,1]^n}$.

**A1** *Start with lists* $G = \emptyset$, $\mathcal{O} = [1]$, *a matrix* $\mathcal{M} = (1,\dots,1)^{\mathrm{tr}} \in \mathrm{Mat}_{s,1}(\mathbb{R})$, *and* $d = 0$.

**A2** *Increase* $d$ *by one and let* $L$ *be the list of all terms of degree* $d$ *in* $\partial\mathcal{O}$, *ordered decreasingly w.r.t.* $\sigma$. *If* $L = \emptyset$, *return the pair* $(G, \mathcal{O})$ *and stop. Otherwise, let* $L = (t_1,\dots,t_\ell)$.

**A3** *Let* $m$ *be the number of columns of* $\mathcal{M}$. *Form the matrix*

$$\mathcal{A} = (\mathrm{eval}(t_1),\dots,\mathrm{eval}(t_\ell),\mathcal{M}) \in \mathrm{Mat}_{s,\ell+m}(\mathbb{R}).$$

*Using its* **SVD**, *calculate a matrix* $\mathcal{B}$ *whose column vectors are an ONB of the* **approximate kernel** $\mathrm{apker}(\mathcal{A},\varepsilon)$.

**A4** Compute the **stabilized reduced row echelon form** of $\mathcal{B}^{\mathrm{tr}}$ with respect to the given $\tau$. The result is a matrix $\mathcal{C} = (c_{ij}) \in \mathrm{Mat}_{k,\ell+m}(\mathbb{R})$ such that $c_{ij} = 0$ for $j < \nu(i)$. Here $\nu(i)$ denotes the column index of the pivot element in the $i^{\mathrm{th}}$ row of $\mathcal{C}$.

**A5** For all $j \in \{1, \ldots, \ell\}$ such that there exists a $i \in \{1, \ldots, k\}$ with $\nu(i) = j$ (i.e. **for the column indices of the pivot elements**), append the polynomial

$$c_{ij} t_j + \textstyle\sum_{j'=j+1}^{\ell} c_{ij'} t_{j'} + \sum_{j'=\ell+1}^{\ell+m} c_{ij'} u_{j'}$$

to the list $G$, where $u_{j'}$ is the $(j' - \ell)^{\mathrm{th}}$ element of $\mathcal{O}$.

**A6** For all $j = \ell, \ell - 1, \ldots, 1$ such that the $j^{\mathrm{th}}$ column of $\mathcal{C}$ contains **no pivot element**, append the term $t_j$ as a new first element to $\mathcal{O}$ and append the column $\mathrm{eval}(t_j)$ as a new first column to $\mathcal{M}$.

**A7** Using the SVD of $\mathcal{M}$, calculate a matrix $\mathcal{B}$ whose column vectors are an ONB of $\operatorname{apker}(\mathcal{M}, \varepsilon)$.

**A8** Repeat steps **A4** – **A7** until $\mathcal{B}$ is empty. Then continue with step **A2**.

This algorithm returns the following results:

**(a)** The set $\mathcal{O} = \{t_1, \ldots, t_\mu\}$ contains an order ideal of terms which is **strongly linearly independent** on $\mathbb{X}$, i.e. such that there is no unitary polynomial in $\langle \mathcal{O} \rangle_K$ which vanishes $\varepsilon$-approximately on $\mathbb{X}$.

**(b)** The set $G$ is a **$\delta$-approximate $\mathcal{O}$-border basis**. (An explicit bound for $\delta$ can be given.)

**Theorem 4.4** **(The ABM Algorithm)**
*Let $\mathbb{X} \subseteq \mathbb{R}^n$ be a finite set of points, and let $\varepsilon > 0$. Consider the following instructions.*

**B1** *Start with lists $G = \emptyset$, $\mathcal{O} = [1]$, a matrix $\mathcal{M} = (1, \ldots, 1)^{\mathrm{tr}} \in \mathrm{Mat}_{s,1}(\mathbb{R})$, and $d = 0$.*

**B2** *Increase $d$ by one and let $L$ be the list of all terms of degree $d$ in $\partial\mathcal{O}$, ordered decreasingly w.r.t. $\sigma$. If $L = \emptyset$, return the pair $(G, \mathcal{O})$ and stop. Otherwise, let $L = (t_1, \ldots, t_\ell)$.*

**B3** *Repeat the following steps for $i = 1, \ldots, \ell$.*

**B4** *Let $A$ be the matrix obtained by appending the column $\mathrm{eval}_{\mathbb{X}}(t_i)$ as a new first column to $M$, let $B = A^{\mathrm{tr}} A$ and $\gamma$ the smallest eigenvalue of $B$.*

**B5** If $\sqrt{\gamma} < \varepsilon$ then compute the norm one eigenvector $(s_0, \ldots, s_m)$ of $B$ w.r.t. $\gamma$ and append $s_0 t_i + s_1 u_1 + \cdots + s_m u_m$ to $G$, where $u_1, \ldots, u_m$ are the elements of $\mathcal{O}$.

**B6** If $\sqrt{\gamma} \geq \varepsilon$ then append $t_i$ to $\mathcal{O}$ and replace $M$ by $A$.

This is algorithm which computes an order ideal $\mathcal{O}$ and an $\delta$-approximate $\mathcal{O}$-border basis $G$ (for some explicitly given $\delta > 0$). The polynomials of $G$ vanish $\varepsilon$-approximately at $\mathbb{X}$, and $\mathcal{O}$ is strongly linearly independent on $\mathbb{X}$.

# 5 – From Approximate to Exact Border Bases

**Impossible only means**
**that you haven't found the solution yet.**
(Anonymous)

**Remark 5.1** In an approximate border basis, the matrices $A_i$ do not commute exactly, but their commutators $A_i\,A_j - A_j\,A_i$ have **small entries**. (The $c_{ij}$ almost satisfy the equations of the border basis scheme.)

**Remark 5.2** Let $\ell = a_1 x_1 + \cdots + a_n x_n$ with $a_i \in \mathbb{R}$, and let $A_1, \ldots, A_n$ be the multiplication matrices of $\mathbb{X}$. Then the points of $\mathbb{X}$ are in normal $\ell$-position if and only if the matrix

$$A_\ell \;=\; a_1\, A_1 + \cdots + a_n\, A_n$$

is **commendable**, i.e. if its characteristic polynomial splits into linear factors and its eigenspaces are 1-dimensional.

**Remark 5.3** If we choose "random" numbers $a_1, \ldots, a_n \in \mathbb{R}$ and let $\ell = a_1 x_1 + \cdots + a_n x_n$ then $\mathbb{X}$ is in normal $\ell$-position and $A_\ell$ is commendable. We shall call $A_\ell$ a **generic multiplication matrix** of $\mathbb{X}$.

The idea is that $A_\ell$ contains all essential information about $\mathbb{X}$. In particular, it determines all $A_i$ and the border basis.

**Proposition 5.4** **(Properties of Multiplication Matrices)**

*Let $\mathbb{X} \subset \mathbb{R}^n$ be a finite point set, and let $\mathcal{O}$ be an order ideal such that there exists an $\mathcal{O}$-border basis of $I_{\mathbb{X}}$.*

**(a)** *The eigenvalues of the $i$-th multiplication matrix $A_i$ w.r.t. $\mathcal{O}$ are the $i$-th coordinates of the points of $\mathbb{X}$.*

**(b)** *Let $\mathcal{O} = \{1, x_1, \ldots, x_n, t_{n+2}, \ldots, t_\mu\}$. Then the joint eigenvectors of $A_1^{\mathrm{tr}}, \ldots, A_n^{\mathrm{tr}}$ are of the form $v_i = (1, p_{i1}, \ldots, p_{in}, \ldots)$ where the points $(p_{i1}, \ldots, p_{in})$ are the points of $\mathbb{X}$.*

## Commuting Families

The characterization of border bases by commuting matrices says that if we find matrices $A_1, \ldots, A_n$ which have the structure of multiplication matrices, they correspond to an actual border basis if and only if they commute.

This suggests the following approach: take the **approximate multiplication matrices** and replace them by near-by rational formal multiplication matrices which commute exactly.

**How do we find such rational matrices?**

**Remark 5.5** If a matrix $B$ is of the form $B = f(A_\ell)$ with a univariate polynomial $f \in \mathbb{Q}[x]$ then $B$ commutes with $A_\ell$.

## The Rational Recovery Algorithm

Suppose that $\widetilde{G}$ is an approximate $\mathcal{O}$-border basis. The following algorithm computes an exact $\mathcal{O}$-border basis **close** to $\widetilde{G}$.

**(R1)** Form the **approximate multiplication matrices** $\widetilde{A}_1, \ldots, \widetilde{A}_n$, i.e. the formal multiplication matrices corresponding to the border prebasis $\widetilde{G}$.

**(R2)** Choose a generic linear form $\ell = a_1 x_1 + \cdots + a_n x_n$ and form the **generic approximate multiplication matrix**

$$\widetilde{A}_\ell = a_1 \, \widetilde{A}_1 + \cdots + a_n \, \widetilde{A}_n$$

**(R3)** Perform **rational recovery** on the matrix $\widetilde{A}_\ell$, i.e. form a rational matrix $A_\ell$ very close to it.

**(4)** Using **Least Squares Approximations**, find polynomials $f_i \in \mathbb{Q}[x]_{\leq \mu}$ such that $f_i(A_\ell)$ is a close to $\widetilde{A}_i$ as possible (for $i = 1, \ldots, n$).

**(5)** Reconstruct the exact border basis $G$ from the commuting matrices $A_i = f_i(A_\ell)$.

**Note:** The fact that the matrices $A_i$ commute follows from the observation that $\mathbb{Q}[A_\ell]$ is a commutative ring.

Let us go carefully through one example for rational recovery.

**Example 5.6** **(The Corners of the Unit Square)**

Let $\mathbb{X} = \{(0.02, 0.01), (0.99, 0.01), (1.01, 0.98), (0.02, 0.99)\}$ be the slightly perturbed corners of the unit square.

We apply the ABM algorithm with $\varepsilon = 0.05$ and get the order ideal $\mathcal{O} = \{1, y, x, xy\}$ and the approximate (unitary) $\mathcal{O}$-border basis $\widetilde{G} = \{\tilde{g}_1, \tilde{g}_2, \tilde{g}_3, \tilde{g}_4\}$ with

$$\begin{aligned}
\tilde{g}_1 &\approx x^2 - 0.02xy - 1.00x + 0.01 \\
\tilde{g}_2 &\approx y^2 - 0.99y \\
\tilde{g}_3 &\approx x^2y - 1.03xy + 0.02y \\
\tilde{g}_4 &\approx xy^2 - 0.98xy
\end{aligned}$$

This yields the approximate multiplication matrices

$$\widetilde{A}_1 = \begin{pmatrix} 0 & 0 & -0.02 & 0 \\ 0 & 0 & 0 & -0.02 \\ 1 & 0 & 1.01 & 0 \\ 0 & 1 & 0.02 & 1.03 \end{pmatrix} \text{ and } \widetilde{A}_2 = \begin{pmatrix} 0 & -0.01 & 0 & 0 \\ 1 & 0.99 & 0 & 0 \\ 0 & -0.01 & 0 & -0.01 \\ 0 & 0 & 1 & 0.99 \end{pmatrix}$$

Both matrices have two eigenvalues very close to 0 and two eigenvalues very close to 1. If we tried to compute the joint eigenvectors of $\widetilde{A}_1^{\mathrm{tr}}$ and $\widetilde{A}_2^{\mathrm{tr}}$, we would get numerically unstable results.

Therefore we form a "random" linear combination $\widetilde{A}_\ell = 0.4\,A_1 + 0.8\,A_2$. Then we consider its (floating point) entries as rational numbers and get $A_\ell$.

Next we approximate $\widetilde{A}_1$ and $\widetilde{A}_1$ as polynomials in $A_\ell$. The result is

$$\begin{array}{rcl}\widetilde{A}_1 & \approx & -0.01\, I_4 + 8.7\, A_\ell - 19.4\, A_\ell^2 + 10.8\, A_\ell^3 \\ \widetilde{A}_2 & \approx & 0.05\, I_4 - 3.1\, A_\ell + 9.7 A_\ell^2 - 5.4 A_\ell^3\end{array}$$

Finally, the right hand sides give truly commuting matrices $A_1$ and $A_2$. From these we read off the exact border basis.

# 6 – An Application: Algebraic Oil

**All models are wrong.**
**Some models are useful.**
(George Box)

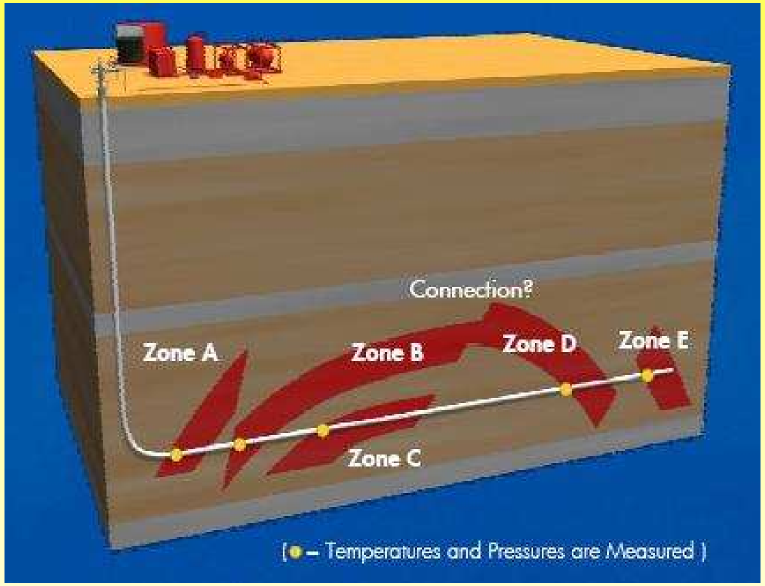
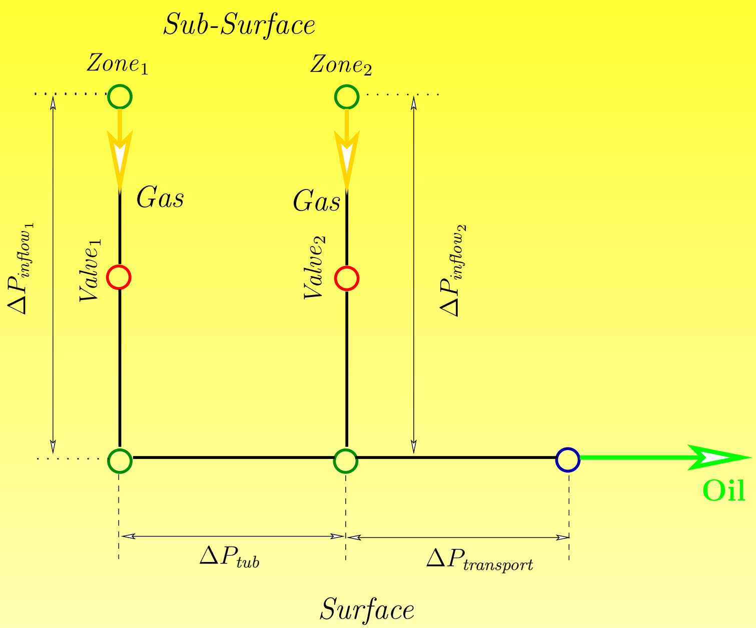



Figure 1: Overview of an Oil Production System

# Schematic Representation of the System




Measurement time series are available for individual zone producing separately **(well test)** and for the situation when they all produce simultaneously.

## The Production Modelling Problem

Assume that **no a priori model** is available to describe the production of a well in terms of measurable physical quantities.

Find a **polynomial model** of the production in terms of the determining, measurable physical quantities which specifically models the **interactions** occurring in this production unit.

Find such a model which correctly **predicts** the behaviour of the production system over **longer time periods** (weeks or even months).

**Example 6.1** A certain two-zone oil well in Brunei yields 7200 data points in $\mathbb{R}^8$. We use 80% of the points for modelling the total production in the following way:

**(a)** Using the ABM algorithm with $\varepsilon = 0.1$, compute an order ideal $\mathcal{O}$ and its evaluation matrix $\mathrm{eval}(\mathcal{O})$.

**(b)** Find the vector in the linear span of the rows of $\mathrm{eval}(\mathcal{O})$ which is closest to the total production.

**(c)** The corresponding linear combination of terms in $\mathcal{O}$ is the **model polynomial** for the total production.

**(d)** Plot all values of the model polynomial at the given points. Compare them at the points which were not used for modelling with the actual measured values.

If you want to know **THE END**,

look at the beginning.

(African Proverb)

**Thank you for your attention!**