Abstract.
Let $K$ be a non-Archimedean valued field. Then, on compact subsets of $K$, every $K$-valued $C^n$-function can be approximated in the $C^n$-topology by polynomial functions (Theorem 1.4). This result is extended to a Weierstrass-Stone type theorem (Theorem 2.10).

INTRODUCTION

The non-archimedean version of the classical Weierstrass Approximation Theorem - the case $n = 0$ of the Abstract - is well known and named after Kaplansky ([1], 5.28). To investigate the case $n = 1$ first let us return to the Archimedean case and consider a real-valued $C^1$-function $f$ on the unit interval. To find a polynomial function $P$ such that both $|f-P|$ and $|f'-P'|$ are smaller or equal than a prescribed $\varepsilon > 0$ one simply can apply the standard Weierstrass Theorem to $f$ obtaining a polynomial function $Q$ for which $|f'-Q| \leq \varepsilon$. Then $x \mapsto P(x) := f(0) + \int_0^x Q(t)dt$ solves the problem.

Now let $f : X \to K$ be a $C^1$-function where $K$ is a non-archimedean valued field and $X \subseteq K$ is compact.

Lacking an indefinite integral the above method no longer works. There do exist continuous linear antiderivations ([3], §64) but they do not map polynomials into polynomials ([3], Ex. 30.C). A further complicating factor is that the natural norm for $C^1$-functions on $X$ is given by

$$
\|f\|_X = \max\{|f(x)| : x \in X\} \lor \max\{\left|\frac{f(x)-f(y)}{x-y}\right| : x, y \in X, x \neq y\}
$$

rather than the more classical formula

$$
\|f\|_X = \max\{|f(x)| : x \in X\} \lor \max\{|f'(x)| : x \in X\}.
$$

(Observe that in the real case both formulas lead to the same norm thanks to the Mean Value Theorem, see [3], §§26,27 for further discussions.)
Thus, to obtain non-archimedean $C^n$-Weierstrass-Stone Theorems for $n \in \{1, 2, \ldots\}$ our methods will necessarily deviate from the 'classical' ones.

0. PRELIMINARIES

1. Throughout $K$ is a non-archimedean complete valued field whose valuation $| \cdot |$ is not trivial. For $a \in K$, $r > 0$ we write $B(a, r) := \{x \in K : |x-a| \leq r\}$, the 'closed' ball about $a$ with radius $r$. 'Clopen' is an abbreviation for 'closed and open'. The function $x \mapsto x$ ($x \in K$) is denoted $X$. The $K$-valued characteristic function of a subset $Y$ of $K$ is written $\xi_Y$. For a set $Z$, a function $f : Z \to K$ and a set $W \subset Z$ we define $\|f\|_W := \sup\{|f(x)| : x \in W\}$ (allowing the value $\infty$). The cardinality of a set $\Gamma$ is $\#\Gamma$. $N_0 := \{0, 1, 2, \ldots\}$, $N := \{1, 2, 3, \ldots\}$.

We now recall some facts from [2], [3] on $C^n$-theory.

2. For a set $Y \subset K$, $n \in N$ we set $\nabla^n Y := \{(y_1, y_2, \ldots, y_n) \in Y^n : i \neq j \Rightarrow y_i \neq y_j\}$. For $f : Y \to K$, $n \in N_0$ we define its $n$th difference quotient $\Phi_n f : \nabla^{n+1} Y \to K$ inductively by $\Phi_0 f := f$ and the formula

$$\Phi_n f(y_1, \ldots, y_{n+1}) = \frac{\Phi_{n-1} f(y_1, y_3, \ldots, y_{n+1}) - \Phi_{n-1} f(y_2, y_3, \ldots, y_{n+1})}{y_1 - y_2}$$

$f$ is called a $C^n$-function if $\Phi_n f$ can be extended to a continuous function on $Y^{n+1}$. The set of all $C^n$-functions $Y \to K$ is denoted $C^n(Y \to K)$. The function $f : Y \to K$ is a $C^\infty$-function if it is in $C^\infty(Y \to K) := \bigcap_{n=0}^\infty C^n(Y \to K)$. The space $C^0(Y \to K)$, consisting of all continuous functions $Y \to K$ is sometimes written as $C(Y \to K)$.

FROM NOW ON IN THIS PAPER $X$ IS A NONEMPTY COMPACT SUBSET OF $K$ WITHOUT ISOLATED POINTS.

3. Since $X$ has no isolated points we have for an $f \in C^n(X \to K)$ that the continuous extension of $\Phi_n f$ to $X^{n+1}$ is unique; we denote this extension by $\overline{\Phi}_n f$. Also we write

$$D_n f(a) := \overline{\Phi}_n f(a, a, \ldots, a) \quad (a \in X)$$

The following facts are proved in [2] and [3].

Proposition 0.3.

(i) For each $n \in N_0$ the space $C^n(X \to K)$ is a $K$-algebra under pointwise operations.
(ii) $C^0(X \to K) \supset C^1(X \to K) \supset \ldots$
(iii) If \( f \in C^n(X \to K) \) then \( f \) is \( n \) times differentiable and \( j!D_jf = f^{(j)} \) for each \( j \in \{0,1,\ldots,n\} \). More generally, if \( i,j \in \{0,1,\ldots,n\} \), \( i+j \leq n \) then \((i+j)!D_iD_jf = D_{i+j}f\).

(iv) If \( f \in C^n(X \to K) \) then for \( x,y \in X \) we have Taylor's formula

\[
f(x) = f(y) + (x-y)D_if(y) + \cdots + (x-y)^{n-1}D_{n-1}f(y) + (x-y)^n \rho_1f(x,y),
\]

where \( \rho_1f(x,y) = \Phi_n f(x,y,y,\ldots,y) \).

4. Since \( X \) is compact the difference quotients \( \Phi_i f \) (\( 0 \leq i \leq n \)) are bounded if \( f \in C^n(X \to K) \). We set

\[
\|f\|_{n,X} := \max\{\|\Phi_i f\|_{n+1,X} : 0 \leq i \leq n\}.
\]

Then \( \|f\|_{0,X} = \|f\|_X \). We quote the following from [2] and [3]. Recall that a function \( f : X \to K \) is a local polynomial if for every \( a \in X \) there is a neighbourhood \( U \) of \( a \) such that \( f | X \cap U \) is a polynomial function.

**Proposition 0.4.** Let \( n \in \mathbb{N}_0 \).

(i) The function \( \| \|_{n,X} \) is a norm on \( C^n(X \to K) \) making it into a \( K \)-Banach algebra.

(ii) The local polynomials form a dense subset of \( C^n(X \to K) \).

(iii) The function

\[
f \mapsto \|f\|_{n,X} := \max_{0 \leq i \leq n-1} \|D_if\|_X \vee \|\rho_1 f\|_{X^2}
\]

(see Proposition 0.3 (iv)) also is a norm on \( C^n(X \to K) \). We have

\[
\|f\|_{n,X} = \max\{\|D_if\|_{n-i,X} : 0 \leq i \leq n\} \quad (f \in C^n(X \to K)).
\]

**Remarks**

1. Proposition 0.4 (ii) will also follow from Proposition 2.8.

2. In general \( \| \|_{n,X} \) is not equivalent to \( \| \|_{n,X} \) for \( n \geq 3 \) (see [3], Example 83.2).
1 THE WEIERSTRASS THEOREM FOR \( C^2 \)-FUNCTIONS

The following product rule for difference quotients is easily proved by induction with respect to \( j \).

Let \( f, g : X \to K \), let \( j \in \mathbb{N}_0 \). Then for all \( (x_1, \ldots, x_{j+1}) \in \nabla^{j+1}X \) we have

\[
\Phi_j(fg)(x_1, \ldots, x_{j+1}) = \sum_{k=0}^{j} \Phi_k f(x_1, \ldots, x_{k+1}) \Phi_{j-k} g(x_{k+1}, \ldots, x_{j+1}).
\]

Or, less precise,

\[
\Phi_j(fg)(x_1, \ldots, x_{j+1}) = \sum_{k=0}^{j} \Phi_k f(x_k) \Phi_{j-k} g(u_{j-k})
\]

for certain \( x_k \in \nabla^{k+1}X \), \( u_{j-k} \in \nabla^{j-k+1}X \).

In the sequel we need an extension of this formula to finite products of functions. The proof is straightforward by induction with respect to \( N \).

**Lemma 1.1. (Product Rule)** Let \( h_1, \ldots, h_N : X \to K \), let \( j \in \mathbb{N}_0 \). Then for all \( (x_1, \ldots, x_{j+1}) \in \nabla^{j+1}X \) we have

\[
\Phi_j(\prod_{s=1}^{N} h_s)(x_1, \ldots, x_{j+1}) = \sum_{s=1}^{N} \prod_{i=1}^{j} \Phi_{i,s} h_s(z_{\sigma,s})
\]

where the sum is taken over all \( \sigma := (j_1, \ldots, j_N) \in \mathbb{N}_0^N \) for which \( j_1 + \cdots + j_N = j \) and where \( z_{\sigma,s} \in \nabla^{j,s+1}X \) for each \( s \in \{1, \ldots, N\} \). (In fact, \( z_{\sigma,1} = (x_1, \ldots, x_{j+1}), z_{\sigma,2} = (x_{j_1+1}, \ldots, x_{j_1+j_2+1}), \ldots, z_{\sigma,N} = (x_{j_1+\cdots+j_N-1+1}, \ldots, x_{j+1}) \). The following key lemma grew out of [1], 5.28.

**Lemma 1.2.** Let \( 0 < \delta < 1 \), \( 0 < \varepsilon < 1 \), let \( B = B_0 \cup B_1 \cup \cdots \cup B_m \) where \( B_0, \ldots, B_m \) are pairwise disjoint 'closed' balls in \( K \) of radius \( \delta \). Then, for each \( n \in \{0, 1, \ldots\} \) there exists a polynomial function \( P : K \to K \) such that \( \|P - \xi_{B_0}\|_{n,B} \leq \varepsilon \).

**Proof.** We may assume \( 0 \in B_0 \). Choose \( c_1 \in B_1, \ldots, c_m \in B_m \); we may assume that \( |c_1| \leq |c_2| \leq \cdots \leq |c_m| \). Then \( \delta < |c_1| \). We shall prove the following statement by induction with respect to \( n \).

Let \( k \in \mathbb{N} \) be such that \( (\delta/|c_1|)^k \leq \varepsilon \delta^n \), \( k > n \). Let \( t_1, t_2, \ldots, t_m \in \mathbb{N} \) be such that for all \( \ell \in \{1, \ldots, m\} \)

\[
|c_\ell|^{k_{t_1}} |c_\ell|^{k_{t_2}} \cdots |c_\ell|^{k_{t_{t-1}}} \left( \frac{\delta}{|c_1|} \right)^t \leq \varepsilon \delta^n
\]
(It is easily seen that such $k, t_1, \ldots, t_m$ exist since $\delta/|c_i| < 1$.) Then the formula

$$P(x) = \prod_{i=1}^{m} (1 - (\frac{x}{c_i})^k)^{t_i}$$

defines a polynomial function $P : K \to K$ for which

$$\|P - \xi_{B_0}\|_{n,B} \leq \varepsilon.$$ 

The case $n = 0$ is proved in [1], 5.28. To prove the step $n - 1 \to n$ we first observe that from the induction hypothesis (with $\varepsilon$ replaced by $\varepsilon\delta$) it follows that

$$\|P - \xi_{B_0}\|_{n-1,B} \leq \varepsilon\delta$$

So it remains to be shown that

$$\|P - \xi_{B_0}\|_{n+1,B} \leq \varepsilon$$

for all $(x_1, \ldots, x_{n+1}) \in \nabla^{n+1}B$. Now, if $|x_i - x_j| > \delta$ for some $i, j \in \{1, \ldots, n+1\}$ we have, using (2),

$$|\Phi_n(P - \xi_{B_0})(x_1, \ldots, x_{n+1})| = |x_i - x_j|^{-1} |\Phi_{n-1}(P - \xi_{B_0})(x_1, x_{j-1}, x_{j+1}, \ldots, x_{n+1}) - \Phi_{n-1}(P - \xi_{B_0})(x_1, x_{i-1}, x_i, x_{i+1}, \ldots, x_{n+1})| \leq \delta^{-1} \varepsilon\delta = \varepsilon.$$ 

So this reduces the proof of (3) to the case where $|x_i - x_j| \leq \delta$ for all $i, j \in \{1, \ldots, n+1\}$; in other words we may assume that $x_1, \ldots, x_{n+1}$ are all in the same $B_\ell$ for some $\ell \in \{0, 1, \ldots, m\}$. But then, after observing that $n \geq 1$, we have $\Phi_n \xi_{B_0}(x_1, \ldots, x_{n+1}) = 0$ so it suffices to prove the following.

If $\ell \in \{0, 1, \ldots, m\}$ and $x_1, \ldots, x_{n+1} \in B_\ell$ are pairwise distinct then

$$|\Phi_n P(x_1, \ldots, x_{n+1})| \leq \varepsilon$$

To prove it we introduce, with $\ell \in \{1, \ldots, m\}$ fixed, the constants $M_i$ ($i \in \{1, \ldots, n\}$) by

$$M_i := \begin{cases} 
1 & \text{if } i > \ell \\
\delta/|c_i| & \text{if } i = \ell \\
|c_i/c_i|^k & \text{if } i < \ell 
\end{cases}$$

and use the following three steps.

**Step 1.** For each $j \in \{0, 1, \ldots, n\}$, $i \in \{1, \ldots, n\}$ we have

$$\|\Phi_j(1 - (\frac{x}{c_i})^k)\|_{\nabla^{j+1}B_\ell} \leq \begin{cases} 
1 & \text{if } \ell = 0, j = 0 \\
\delta^{-j}(\frac{\delta}{|c_i|})^k & \text{if } \ell = 0, j > 0 \\
\delta^{-j}M_i & \text{if } \ell > 0.
\end{cases}$$
Proof.

a. The case \( j = 0 \). Then for \( x \in B_\ell \) we have
- if \( i > \ell \) then \( |1 - \left( \frac{x}{c_i} \right)^k| = 1 \)
- if \( i = \ell \) then \( |1 - \left( \frac{x}{c_i} \right)^k| = \left| \frac{c_i - x}{c_i} \right|^k \leq \frac{\delta^k}{|c_i|^k} \leq \frac{\delta}{|c_i|} \)
- if \( i < \ell \) then \( |1 - \left( \frac{x}{c_i} \right)^k| = \left| \frac{x}{c_i} \right|^k \leq \frac{|c_i|}{|c_i|^k} \)

and the statement follows.

b. The case \( j > 0 \). Then \( \Phi_j(1) = 0 \) so that

\[
\Phi_j(1 - \left( \frac{x}{c_i} \right)^k) = \frac{1}{c_i^k} \Phi_j(x^k)
\]

Let \( (x_1, \ldots, x_{j+1}) \in \nabla^{j+1}B_\ell \). By the Product Rule 1.1, \( \Phi_j(x^k)(x_1, \ldots, x_{j+1}) \) is a sum of terms of the form \( \prod_{s=1}^{k} (\Phi_j(x^k)(z_s)) \). Such a term is 0 if one of the \( j_s \) is > 1, so we only have to deal with \( j_s = 0 \) (then \( \Phi_j(x^k)(z_s) = x^k \) or \( j_s = 1 \) (then \( \Phi_j(x^k)(z_s) = 1 \)). The latter case occurs \( j \) times (as \( \sum j_s = j \)) and it follows that

\[
\prod_{s=1}^{k} (\Phi_j(x^k)(z_s)) \]

is a product of \( k-j \) distinct terms taken from \( \{x_1, \ldots, x_{j+1}\} \) (observe that, indeed, \( j < k \) since \( j \leq n < k \)), so its absolute value is \( \leq |c_i|^{k-j} \). It follows that \( \|\Phi_j(1 - (\frac{x}{c_i})^k)\|_{\nabla^{j+1}B_\ell} \leq |c_i|^{k-j}/|c_i|^k \) from which we conclude
- if \( \ell = 0 \) : \( |c_i|^{k-j}/|c_i|^k \leq \delta^{k-j}/|c_i|^k = \delta^{-j}(|c_i|)^k \),
- if \( i > \ell > 0 \) : \( |c_i|^{k-j}/|c_i|^k \leq |c_i|^{k-j} \leq \delta^{-j} = \delta^{-j} M_i \)
- if \( i = \ell > 0 \) : \( |c_i|^{k-j}/|c_i|^k \leq |c_i|^{k-j} \leq |c_i|^{k-j} \leq \delta^{-j} M_i \)
- if \( i < \ell \) : \( |c_i|^{k-j}/|c_i|^k \leq |c_i|^{k-j} \leq \delta^{-j} M_i \)

and step 1 is proved.

Step 2. For each \( j \in \{0, 1, \ldots, n\}, i \in \{1, \ldots, n\} \) we have

\[
\|\Phi_j(1 - (\frac{x}{c_i})^k)\|_{\nabla^{j+1}B_\ell} \leq \begin{cases} 
1 & \text{if } \ell = 0, j = 0 \\
\delta^{-j}(|c_i|)^k & \text{if } \ell = 0, j > 0 \\
\delta^{-j} M_i^k & \text{if } \ell > 0
\end{cases}
\]

Proof. The case \( j = 0 \) follows directly from Step 1, part a, so assume \( j > 0 \). By the Product Rule 1.1 applied to \( h_s = 1 - (\frac{x}{c_i})^k \) for all \( s \in \{1, \ldots, t_i\} \) we have for \( (x_1, \ldots, x_{j+1}) \in \nabla^{j+1}B_\ell \) that \( \Phi_j(1 - (\frac{x}{c_i})^k)(x_1, \ldots, x_{j+1}) \) is a sum of terms of the form

\[
\prod_{s=1}^{t_i} \Phi_j(1 - (\frac{x}{c_i})^k)(z_s)
\]
where \( j_1 + \cdots + j_s = j \). If \( \ell = 0 \) it follows from Step 1 that the absolute value of (5) is \( \leq \prod \delta^{-j_s}(\frac{\delta}{|c_i|})^k \) where the product is taken over all \( s \) in the nonempty set \( \Gamma := \{ s \in \{1, \ldots, t_i\} : j_s > 0 \} \), so the product is \( \leq \delta^{-j}(\frac{\delta}{|c_i|})^{k} \leq \delta^{-j}(\frac{\delta}{|c_i|})^{k} \). If \( \ell > 0 \) it follows from Step 1 that the absolute value of (5) is \( \leq \prod \delta^{-j_s} M_i = \delta^{-j} M_i^t \).

The statement of Step 2 follows.

**Step 3.** Proof of (4). Again, the Product Rule 1.1, now applied to \( h_i = (1 - (\frac{x}{c_i})^k)^t \) for \( i \in \{1, \ldots, m\} \) tells us that for \( (x_1, \ldots, x_{n+1}) \in \nabla^{n+1} B_\ell \) the expression \( \Phi_n P(x_1, \ldots, x_{n+1}) \) is a sum of terms of the form

\[
(6) \quad \prod_{i=1}^m \Phi_n_i (1 - (\frac{x}{c_i})^k)^t (z_s)
\]

where \( n_1 + \cdots + n_m = n \). If \( \ell = 0 \) we have by Step 2 that the absolute value of (6) is \( \leq \prod \delta^{-n_i}(\frac{\delta}{|c_i|})^{k_i} \) where the product is taken over \( i \) in the nonempty set \( \Gamma := \{ i : n_i \neq 0 \} \), so the product is \( \leq \delta^{-n}(\frac{\delta}{|c_i|})^{k} \leq \delta^{-n} \cdot \epsilon \delta^n = \epsilon \), where we used the assumption \( (\delta/|c_i|)^k \leq \epsilon \delta^n \). We see that \( \Phi_n P(x_1, \ldots, x_{n+1}) \leq \epsilon \) if \( (x_1, \ldots, x_n) \in B_\ell \) now let \( \ell > 0 \). By Step 2 we have that the absolute value of (6) is \( \leq \prod_{i=1}^m \delta^{-n_i} M_i^t = \delta^{-n} M_1^t \cdots M_m^t = \delta^{-n} \cdot \frac{\delta^{k_1} \cdots \delta^{k_m}}{|c_1|} \cdot \frac{\delta^{k_1} \cdots \delta^{k_m}}{|c_m|} \) which is \( \leq \delta^{-n} \epsilon \delta^n \) by (1). This proves (4) and the Lemma.

**Corollary 1.3.** For every locally constant \( f : X \rightarrow K \), for every \( n \in \mathbb{N}_0 \) and \( \epsilon > 0 \) there exists a polynomial function \( P : K \rightarrow K \) such that \( \| f - P \|_{n,X} \leq \epsilon \).

**Proof.** There exist a \( \delta \in (0,1) \), pairwise disjoint 'closed' balls \( B_1, \ldots, B_m \) of radius \( \delta \) covering \( X \) and \( \lambda_1, \ldots, \lambda_m \in K \) such that

\[ f(x) = \sum_{i=1}^m \lambda_i \xi B_i(x) \quad (x \in X) \]

By Lemma 1.2 there exist polynomials \( P_1, \ldots, P_m \) such that \( \| \xi B_i - P_i \|_{n,X} \leq \| \xi B_i - P_i \|_{n,\cup B_i} \leq \epsilon (|\lambda_i| + 1)^{-1} \) for each \( i \in \{1, \ldots, m\} \). Then \( P := \sum_{i=1}^m \lambda_i P_i \) is a polynomial function and \( \| f - P \|_{n,X} \leq \max_i \| \lambda_i (\xi B_i - P_i) \|_{n,X} \leq \max_i |\lambda_i| \epsilon (|\lambda_i| + 1)^{-1} \leq \epsilon \).

**Theorem 1.4. (\( C^n \)-Weierstrass Theorem)** For each \( n \in \mathbb{N}_0 \), \( f \in C^n(X \rightarrow K) \) and \( \epsilon > 0 \) there exists a polynomial function \( P : K \rightarrow K \) such that \( \| f - P \|_{n,X} \leq \epsilon \).

**Proof.** There is by Proposition 0.4 a local polynomial \( g : K \rightarrow K \) with \( \| f - g \|_{n,X} \leq \epsilon \). This \( g \) has the form \( g = \sum_{i=1}^m Q_i h_i \) where \( Q_1, \ldots, Q_m \) are polynomials and \( h_1, \ldots, h_m \)
are locally constant. By Corollary 1.3 we can find polynomials $P_1, \ldots, P_m$ for which $\|h_i - P_i\|_{n, X} \leq \epsilon (\|Q_i\|_{n, X} + 1)^{-1}$ for each $i$. Then $P := \sum_{i=1}^m Q_i P_i$ is a polynomial and $\|g - P\|_{n, X} \leq \epsilon$. It follows that $\|f - P\|_{n, X} \leq \max(\|f\|_{n, X}, \|g - P\|_{n, X}) \leq \epsilon$.

Remarks.
1. In the case where $X = \mathbb{Z}_p$, $K \supset \mathbb{Q}_p$ the above Theorem 1.4 is not new: The Mahler base $e_0, e_1, \ldots$ of $C(\mathbb{Z}_p \to K)$ defined by $e_m(x) = (\frac{2}{m})$ is proved in [3], §54 to be a Schauder base for $C^n(\mathbb{Z}_p \to K)$, for each $n$.
2. It follows directly from Theorem 1.4 that the polynomial functions $X \to K$ form a dense subset of $C^\infty(X \to K)$.

2. A WEIERSTRASS-STONE THEOREM FOR $C^n$-FUNCTIONS

For this Theorem (2.10) we will need the continuity of $g \mapsto g \circ f$ in the $C^n$-topologies (Proposition 2.5). To prove it we need some technical lemmas that are in the spirit of [3], §77.

Let $n \in \mathbb{N}$. For a function $h : \nabla^n X \to K$ we define $\Delta h : \nabla^{n+1} X \to K$ by the formula

$$\Delta h(x_1, x_2, \ldots, x_{n+1}) = \frac{h(x_1, x_3, x_4, \ldots, x_{n+1}) - h(x_2, x_3, \ldots, x_{n+1})}{x_1 - x_2}$$

We have the following product rule.

Lemma 2.1. (Product Rule). Let $n \in \mathbb{N}$, let $h, t : \nabla^n X \to K$. Then for all $(x_1, x_2, \ldots, x_{n+1}) \in \nabla^{n+1} X$ we have $\Delta(ht)(x_1, x_2, \ldots, x_{n+1}) = h(x_2, x_3, \ldots, x_{n+1}) \Delta t(x_1, x_2, \ldots, x_{n+1}) + t(x_1, x_3, \ldots, x_{n+1}) \Delta h(x_1, x_2, \ldots, x_{n+1})$.

Proof. Straightforward.

Lemma 2.2. Let $f : X \to K$, $n \in \mathbb{N}_0$. Let $S_n$ be the set of the following functions defined on $\nabla^{n+1} X$.

$$(x_1, \ldots, x_{n+1}) \mapsto \Phi_1 f(x_{i_1}, x_{i_2}) \quad (1 \leq i_1 < i_2 \leq n + 1)$$

$$(x_1, \ldots, x_{n+1}) \mapsto \Phi_2 f(x_{i_1}, x_{i_2}, x_{i_3}) \quad (1 \leq i_1 < i_2 < i_3 \leq n + 1)$$

$$\vdots$$

$$(x_1, \ldots, x_{n+1}) \mapsto \Phi_n f(x_1, \ldots, x_{n+1}).$$

For $k \in \mathbb{N}$, let $R_n^k$ be the additive group generated by $S_n, S_n^2, \ldots, S_n^k$ where, for each $j \in \{1, \ldots, k\}$, $S_n^j$ is the product set $\{h_1 h_2 \ldots h_j : h_i \in S_n \text{ for each } i \in \{1, \ldots, j\}\}$. Then, for all $k, n \in \mathbb{N}$, $\Delta R_n^k \subset R_n^{k+1}$. 
**Proof.** We use induction with respect to $k$. For the case $k = 1$ it suffices to prove $h \in S_n \Rightarrow \Delta h \in R^1_{n+1}$. Then $h$ has the form

$$(x_1, \ldots, x_{n+1}) \mapsto \Phi_j f(x_{i_1}, x_{i_2}, \ldots, x_{i_{j+1}})$$

for some $j \in \{2, 3, \ldots, n+1\}$ and so

$$\Delta h(x_1, x_2, \ldots, x_{n+1}) = \frac{h(x_1, x_3, \ldots, x_{n+2}) - h(x_2, x_3, \ldots, x_{n+2})}{x_1 - x_2}$$

vanishes if $i_1 > 1$ (and then $\Delta h$ is the null function), while if $i_1 = 1$ it equals

$$\frac{\Phi_j f(x_1, x_{i+1}, \ldots, x_{i+1}) - \Phi_j f(x_2, x_{i+1}, \ldots, x_{i+1})}{x_1 - x_2} = \Phi_{j+1} f(x_1, x_2, x_{i+2}, \ldots, x_{i+1})$$

and it follows that $\Delta h \in S_{n+1} \subset R^1_{n+1}$. For the induction step assume $\Delta R^k_{n+1} \subset R^k_{n+1}$; it suffices to prove that $\Delta S^k_n \subset R^k_{n+1}$. So let $h \in S^k_n$ and write $h = h_1 H$, where $h_1 \in S_n$, $H \in S^k_n$. By the Product Rule 2.1 we have

$$\Delta h(x_1, \ldots, x_{n+2}) = h_1(x_2, x_3, \ldots, x_{n+2}) \Delta H(x_1, x_2, \ldots, x_{n+2}) +$$

$$+ H(x_1, x_3, \ldots, x_{n+2}) \Delta h_1(x_1, x_2, \ldots, x_{n+2}).$$

The fact that $h_1 \in S_n$ makes

$$(x_1, x_2, \ldots, x_{n+2}) \mapsto h_1(x_1, x_3, \ldots, x_{n+2})$$

into an element of $S_{n+1}$. Similarly, since $H \in S^k_{n+1}$, the function

$$(x_1, x_2, \ldots, x_{n+2}) \mapsto H(x_2, x_3, \ldots, x_{n+2})$$

is in $S^k_{n+1}$. By our first induction step, $\Delta h_1 \in R^1_{n+1}$ and by the induction hypothesis $\Delta H \in R^k_{n+1}$. Hence,

$$\Delta h \in S_{n+1} R^k_{n+1} + S^k_{n+1} R^1_{n+1} \subset R^1_{n+1} R^k_{n+1} + R^1_{n+1} R^k_{n+1} \subset R^k_{n+1}.$$

**Lemma 2.3.** Let $f, n, S_n, k, R^k_n$ be as in the previous lemma. Let $f(X) \subset Y \subset K$ where $Y$ has no isolated points. Let $g : Y \rightarrow K$ be a $C^n$-function. Let $B_n$ be the set of the following functions defined on $\nabla^{n+1} X$.

$$(x_1, \ldots, x_{n+1}) \mapsto \Phi_1 g(f(x_{i_1}), f(x_{i_2})) \quad (1 \leq i_1 < i_2 \leq n + 1)$$

$$(x_1, \ldots, x_{n+1}) \mapsto \Phi_2 g(f(x_{i_1}), f(x_{i_2}), f(x_{i_3})) \quad (1 \leq i_1 < i_2 < i_3 \leq n + 1)$$

$$\vdots$$

$$(x_1, \ldots, x_{n+1}) \mapsto \Phi_n g(f(x_1), f(x_2), \ldots, f(x_{n+1})).$$
Let $A_n$ be the additive group generated by $B_nR^n$. Then

$$\Delta A_n \subseteq A_{n+1}.$$ 

**Proof.** We prove: $h \in B_nR^n \Rightarrow \Delta h \in A_{n+1}$. Write $h = br$ where $b \in B_n$, $r \in R^n$. By the Product Rule 2.1 we have for all $(x_1, x_2, \ldots, x_{n+2}) \in \nabla^{n+2}X$

$$\Delta h(x_1, x_2, \ldots, x_{n+2}) = b(x_2, x_3, \ldots, x_{n+2}) \Delta r(x_1, x_2, \ldots, x_{n+2}) +$$

$$+ r(x_1, x_3, \ldots, x_{n+2}) \Delta b(x_1, x_2, \ldots, x_{n+2}).$$

We have:

(i) $b \in B_n$ so $(x_1, \ldots, x_{n+2}) \mapsto b(x_2, x_3, \ldots, x_{n+1})$ is in $B_{n+1}$.

(ii) $r \in R^n$ so $(x_1, \ldots, x_{n+2}) \mapsto r(x_1, x_3, \ldots, x_{n+1})$ is in $R_{n+1}^n$ (in the previous proof we had $r \in S^k_n \Rightarrow$ the map $(x_1, \ldots, x_{n+2}) \mapsto r(x_1, x_3, \ldots, x_{n+1})$ is in $S^k_{n+1}$, and (ii) follows from this).

(iii) $r \in R^n$ so $\Delta r \in R_{n+1}^n$ (Previous Lemma).

(iv) $b$ has the form

$$(x_1, x_2, \ldots, x_{n+1}) \mapsto \overline{\Phi}_j g(f(x_{i_1}), \ldots, f(x_{i_{j+1}}))$$

for some $j \in \{2, \ldots, n+1\}$ and so

$$\Delta b(x_1, x_2, \ldots, x_{n+2}) = \frac{b(x_1, x_3, x_4, \ldots, x_{n+2}) - b(x_2, x_3, \ldots, x_{n+2})}{x_1 - x_2}$$

vanishes if $i_1 > 1$ (and then $\Delta b$ is the null function), while if $i_1 = 1$ it equals

$$\overline{\Phi}_j g(f(x_1), f(x_{i_1+1}), \ldots, f(x_{i_{j+1}+1})) - \overline{\Phi}_j g(f(x_2), f(x_{i_1+1}), \ldots, f(x_{i_{j+1}+1}))$$

$$= \overline{\Phi}_{j+1} g(f(x_1), f(x_2), f(x_{i_2+1}), \ldots, f(x_{i_{j+1}+1})) \Phi_1 f(x_1, x_2).$$

(if $f(x_1) = f(x_2)$ we have 0 at both sides). So we see that $\Delta b \in B_{n+1}R_{n+1}^n$.

Combining (i) - (iv) we get $\Delta h \in B_{n+1}R_{n+1}^n + R_{n+1}^n B_{n+1}R_{n+1}^n \subseteq B_{n+1}R_{n+1}^{n+1} + B_{n+1} \cdot R_{n+1}^{n+1} \subseteq A_{n+1}$.

**Corollary 2.4.** With the notations as in the previous lemma we have $\Phi_n(g \circ f) \in A_n$ ($n \in \mathbb{N}$).

**Proof.** We proceed by induction on $n$. For the case $n = 1$ we write, for $(x_1, x_2) \in \nabla^2 X$,

$$\Phi_1(g \circ f)(x_1, x_2) = (x_1 - x_2)^{-1} \left( g(f(x_1)) - g(f(x_2)) \right) = \overline{\Phi}_1 g(f(x_1), f(x_2)) \Phi_1 f(x_1, x_2).$$
Hence, $\Phi_1(g \circ f) \in B_1 S_1 \subset B_1 R \subset A_1$. To prove the step $n \rightarrow n+1$ observe that by the induction hypothesis, $\Phi_n(g \circ f) \in A_n$. By Lemma 2.3, $\Phi_{n+1}(g \circ f) = \Delta \Phi_n(g \circ f) \in A_{n+1}$.

**Remark.** From Corollary 2.4 it follows easily that the composition of two $C^n$-functions is again a $C^n$-function, a result that already was obtained in [3], 77.5.

**Proposition 2.5.** (Continuity of $g \mapsto g \circ f$) Let $n \in N_0$, let $f \in C^n(X \rightarrow K)$ and let $g \in C^n(Y \rightarrow K)$ where $Y$ has no isolated points, $Y \supset f(X)$. Then $\|g \circ f\|_{n,X} \leq \|g\|_{n,Y} \max_{0 \leq j \leq n} \|f\|_{j,X}$.

**Proof.** We may assume $\|g\|_{n,Y} < \infty$. It suffices to prove $\|\Phi_n(g \circ f)\|_{n,X} \leq \|g\|_{n,Y} \|f\|_{n,X}$. Now $\|\Phi_0(g \circ f)\|_{n,X} = \max_{x \in X} |g(f(x))| \leq \|g\|_{0,Y} \|f\|_{0,X}$ which proves the case $n = 0$. For $n \geq 1$ we apply Corollary 2.4 which says that $\Phi_n(g \circ f) \in A_n$ i.e. $\Phi_n(g \circ f)$ is a sum of functions in $B_n S_n^n$. By the definition of $B_n$ we have

\[ (*) \quad h \in B_n \Rightarrow \|h\|_{n,X} \leq \|g\|_{n,Y} \]

Similarly

\[ k \in S_n \Rightarrow \|k\|_{n,X} \leq \max_{1 \leq i \leq n} \|\Phi_i f\|_{n,X} \leq \|f\|_{n,X} \]

so that

\[ (**) \quad k \in S_n \Rightarrow \|k\|_{n,X} \leq \|f\|_{n,X}. \]

Combination of $(*)$ and $(**)$ yields $\|\Phi_n(g \circ f)\|_{n,X} \leq \|g\|_{n,Y} \|f\|_{n,X}$.

Proposition 2.5 enables us to prove

**Proposition 2.6.** Let $n \in N_0$ and let $A$ be a closed subalgebra of $C^n(X \rightarrow K)$. Suppose $A$ separates the points of $X$ and contains the constant functions. Then $A$ contains all locally constant functions $X \rightarrow K$.

**Proof.** 1. We first prove that $f \in A$, $U \subset K$, $U$ clopen implies $\xi_{f^{-1}(U)} \in A$. In fact, $f(X)$ is compact so there exist a $\delta \in (0,1)$ and finitely many disjoint balls $B_1, \ldots, B_m$ of radius $\delta$ covering $f(X)$ where, say, $B_1 \ldots B_q$ lie in $U$, and $B_{q+1}, \ldots, B_m$ are in $K \setminus U$. Let $\varepsilon > 0$. By the Key Lemma 1.2 there exists, for each $i \in \{1, \ldots, m\}$ a polynomial $P_i$ such that $\|\xi_{B_i} - P_i\|_{n,B} < \varepsilon$, where $B := \bigcup_{i=1}^m B_i$. Then $P := \sum_{i=1}^q P_i$ is a polynomial and

\[ \|P - \xi_U\|_{n,B} = \|P - \xi_{B^0}\|_{n,B} = \| \sum_{i=1}^q (P_i - \xi_{B_i}) \|_{n,B} < \varepsilon, \]

where $B^0 := \bigcup_{i=1}^q B_i$.

By Proposition 2.5

\[ \|(P - \xi_U) \circ f\|_{n,X} \leq \|P - \xi_U\|_{n,B} \max_{0 \leq j \leq n} \|f\|_{j,X} \leq \varepsilon \max_{0 \leq j \leq n} \|f\|_{j,X} \]
and we see that there exists a sequence $P_1, P_2, \ldots$ of polynomials such that
\[ \|P_k \circ f - \xi U \circ f\|_{n, X} \to 0. \]
Since $A$ is an algebra with an indentity we have $P_k \circ f \in A$ for all $k$. Then $\xi f^{-1}(U) = \xi U \circ f$ = \lim_{k \to \infty} P_k \circ f \in A$.

2. Now consider
\[ B := \{ V \subset X, \xi V \in A \}. \]
It is very easy to see that $B$ is a ring of clopen subsets of $X$ and that $B$ covers $X$. To show that $B$ separates the points of $X$ let $x \in X, y \in X, x \neq y$. Then there is an $f \in A$ for which $f(x) \neq f(y)$. Set $U := \{ \lambda \in K : |\lambda - f(x)| < |f(x) - f(y)| \}$. Then $U$ is clopen in $K$. By the first part of the proof, $f^{-1}(U) \in B$. But $x \in f^{-1}(U)$ whereas $y \notin f^{-1}(U)$.

By [1], Exercise 2.8 $B$ is the ring of all clopens of $X$. It follows easily that all locally constant functions are in $A$.

To arrive at the Weierstrass-Stone Theorem 2.10 we need a final technical lemma.

**Lemma 2.7.** Let $a_1, \ldots, a_m \in X$, let $\delta_1, \ldots, \delta_m$ be in $(0,1)$ such that $B(a_1, \delta_1), \ldots, B(a_m, \delta_m)$ form a disjoint covering of $X$. Let $n \in \mathbb{N}_0$, $h \in C^n(X \to K)$ and suppose $D_j h(a_i) = 0$ and $|\overline{f}_{n-j} D_j h(x_1, \ldots, x_{n-j+1})| \leq \varepsilon$ for all $i \in \{1, \ldots, m\}$, $x_1, \ldots, x_{n+1} \in B(a_i, \delta_i) \cap X$, $j \in \{0, 1, \ldots, n\}$. Then $\|h\|_{n, X} \leq \varepsilon$.

**Proof.** We first prove that $\|h\|_{n, X} \leq \varepsilon$ (see Proposition 0.4(iii)). Let $i \in \{1, \ldots, m\}$. Set $B_i = B(a_i, \delta_i)$. By Taylor's formula (Proposition 0.3(iv)) we have for $x \in X \cap B_i$:
\[ |h(x)| = \left| \sum_{j=0}^{n-1} (x - a_i)^j D_j h(a_i) + (x - a_i)^n \rho_1 h(x, a_i) \right| \leq \delta_i^n \varepsilon. \]
Similarly we have for $j \in \{0, \ldots, n - 1\}$ and $x \in X \cap B_i : |D_j h(x)| = \left| \sum_{j=0}^{n-1-j} (x - a_i)^j D_{j+1} h(a_i) + (x - a_i)^{n-j} \rho_1 (D_j h)(x, a_i) \right|$. Now using Proposition 0.3(iii) we see that $D_j D_j h(a_i) = 0$ so that
\[ (*) \quad |D_j h(x)| = |x - a_i|^{n-j} |\overline{f}_{n-j} D_j h(x, a_i, \ldots, a_i)| \leq \delta_i^{n-j} \varepsilon. \]
It follows that $\|h\|_X, \|D_1 h\|_X, \ldots, \|D_{n-1} h\|_X$ are all $\leq \varepsilon$. Now let $x, y \in X$. If $x, y$ are in the same $B_i$ then $|\rho_1 h(x, y)| = |\overline{f}_{n} h(x, y, \ldots, y)| \leq \varepsilon$ by assumption. If $x \in B_i, y \in B_j$ and $i \neq j$ then $|x - y| \geq \delta := \max(\delta_i, \delta_j)$ and by Taylor's formula
\[ h(x) = \sum_{i=0}^{n-1} (x - y)^i D_i h(y) + (x - y)^n \rho_1 h(x, y) \]
we obtain, using $(*)$,
\[ |\rho_1 h(x, y)| \leq \frac{|h(x) - h(y)|}{|(x - y)^n|} \vee \frac{|D_1 h(y)|}{|x - y|^{n-1}} \vee \ldots \vee \frac{|D_{n-1} h(y)|}{|x - y|} \leq \frac{\delta^n \varepsilon}{\delta^n} \vee \frac{\delta^{n-1} \varepsilon}{\delta^{n-1}} \vee \ldots \vee \frac{\delta \varepsilon}{\delta} \leq \varepsilon. \]
and we have proved \( \|h\|_{n,X} \leq \varepsilon \).

Now to prove that even \( \|h\|_{n,X} \leq \varepsilon \) observe that by Proposition 0.4(iii)

\[
\|h\|_{n,X} = \|h\|_{n-1,X} \lor D_1 h \lor \cdots \lor D_n h.
\]

To prove, for example, that \( \|D_1 h\|_{n-1,X} \leq \varepsilon \) we observe that by Proposition 0.4(iii) \( D_1 h \in C^{n-1}(X \to K) \) and that for \( i \in \{0, \ldots, n-2\} \) and \( j \in \{0, 1, \ldots, n-2\} \) we have \( D_j D_1 h(a_i) = (j+1)D_{j+1} h(a_i) = 0 \) and for all \( x_1, \ldots, x_n \in B(a_i, \delta_i) \) and \( j \in \{0, 1, \ldots, n-2\} \)

\[
[\mathcal{F}_{n-1-j} D_j(D_1 h)(x_1, \ldots, x_{n-j})] = [(j + 1)] [\mathcal{F}_{n-1-j} D_{j+1} h(x_1, \ldots, x_{n-j})] \leq \varepsilon
\]

by assumption. So the conditions of our Lemma (with \( D_1 h, n-1 \) in place of \( h, n \) respectively) are satisfied and by the first part of the proof we may conclude that \( \|D_1 h\|_{n-1,X} \leq \varepsilon \). In a similar way we prove that \( \|D_2 h\|_{n-2,X} \leq \varepsilon, \ldots, \|D_n h\|_{0,X} \leq \varepsilon \) and it follows that \( \|h\|_{n,X} \leq \varepsilon \).

**Proposition 2.8.** Let \( n \in N_0 \) and let \( A \) be a closed subalgebra of \( C^n(X \to K) \) containing the locally constant functions. Let \( g \in C^n(X \to K) \) and suppose for each \( a \in X \) there exists an \( f_a \in A \) with \( D_i g(a) = D_i f_a(a) \) for \( i \in \{0, 1, \ldots, n\} \). Then \( g \in A \).

**Proof.** Let \( \varepsilon > 0 \). For each \( a \in X \) choose an \( f_a \in A \) with \( f_a(a) = g(a) \), \( D_1 f_a(a) = D_1 g(a) \), \ldots, \( D_n f_a(a) = D_n g(a) \). By continuity there exists a \( \delta_a > 0 \) such that, with \( h_a := f_a - g, [\mathcal{F}_{n-1-j} D_j h_a(x_1, \ldots, x_{n-j+1})] \leq \varepsilon \) for all \( j \in \{0, 1, \ldots, n\} \) and \( x_1, \ldots, x_{n-j+1} \in B(a, \delta_a) \). The \( B(a, \delta_a) \) cover \( X \) and by compactness there exists a finite disjoint subcovering \( B(a_1, \delta_{a_1}), \ldots, B(a_m, \delta_{a_m}) \). Set

\[
f := \sum_{i=1}^m f_{a_i} \chi_{B(a_i, \delta_{a_i})} \cap X
\]

Then, by our assumption on \( A, f \in A \). By Lemma 2.7, applied to \( h := f - g \) and where \( \delta_1, \ldots, \delta_m \) are replaced by \( \delta_{a_1}, \ldots, \delta_{a_m} \) respectively, we then have \( \|f - g\|_{n,X} \leq \varepsilon \). We see that \( g \in \overline{A} = A \).

**Remark.** It follows directly that the local polynomial functions \( X \to K \) form a dense subset of \( C^n(X \to K) \).

**Proposition 2.9.** Let \( n \in N \) and let \( A \) be a \( K \)-subalgebra of \( C^n(X \to K) \) containing the constant functions. Suppose \( f'(a) \neq 0 \) for some \( f \in A, a \in X \). Then there is a \( g \in A \) with \( g(a) = 0, g'(a) = 1 \) and \( D_2 g(a) = D_3 g(a) = \cdots = D_n g(a) = 0 \).

**Proof.** By considering the function \( f'(a)^{-1}(f - f(a)) \) it follows that we may assume that \( f(a) = 0, f'(a) = 1 \). Then

\[
f = (X - a) h
\]
where $h$ is continuous, $h(a) = 1$. To obtain the statement by induction with respect to $n$ we only have to consider the induction step $n - 1 \to n$ and, to prove that, we may assume that $D_2 f(a) = \cdots = D_{n-1} f(a) = 0$. From (*) we obtain
\[
 f^n = (X - a)^n h^n
\]
and by uniqueness of the Taylor expansion of the $C^n$-function $f^n$ we obtain $f^n(a) = D_1 f^n(a) = \cdots = D_{n-1} f^n(a) = 0$ and $D_n f^n(a) = h^n(a) = 1$. We see that $g := f - D_n f(a) f^n$ is in $A$ and that $g(a) = 0$, $g'(a) = 1$, $D_2 g(a) = \cdots = D_{n-1} g(a) = 0$ and $D_n g(a) = D_n f(a) - D_n f(a) D_n f^n(a) = 0$.

**Theorem 2.10. (Weierstrass-Stone Theorem for $C^n$-functions).** Let $n \in \mathbb{N}$ and let $A$ be a closed subalgebra that separates the points of $X$ and that contains the constant functions. Suppose also that for each $a \in X$ there exists an $f \in A$ with $f'(a) \neq 0$. Then $A = C^n(X \to K)$.

**Proof.** By Proposition 2.9, for each $a \in X$ there exists an $f \in A$ with $f(a) = 0$, $f'(a) = 1$, $D_i f(a) = 0$ for $i \in \{2, \ldots, n\}$. The function $g := X$ satisfies $g(a) = 0$, $g'(a) = 1$, $D_i g(a) = 0$ for $i \in \{2, \ldots, n\}$ so applying Proposition 2.8 (observe that $A$ contains the locally constant functions by Proposition 2.6) we obtain that $X \in A$. But then all polynomials are in $A$ and $A = C^n(X \to K)$ by the Weierstrass Theorem 1.4.

**Remarks.**

1. The case $n = 0$ yields, at least for those $X$ that are embeddable into $K$, the well known Kaplansky Theorem proved in [1], 6.15.

2. We leave it to the reader to establish a $C^\infty$-version of Theorem 2.10.

**REFERENCES**

