A CHEBYSHEV-TYPE ALTERNATION THEOREM FOR BEST APPROXIMATION BY A SUM OF TWO ALGEBRAS

AIDA KH. ASGAROVA¹, ALI A. HUSEYNLI¹,² AND VUGAR E. ISMAILOV¹

¹Department of Function Theory, Institute of Mathematics and Mechanics, Baku, AZ1141, Azerbaijan (aidaasgarova@gmail.com; aidaasgarova@gmail.com)
²Department of Mathematics, Khazar University, Baku, AZ1096, Azerbaijan (alihuseynli@gmail.com)

(Received 8 December 2022)

Abstract Let $X$ be a compact metric space, $C(X)$ be the space of continuous real-valued functions on $X$ and $A_1, A_2$ be two closed subalgebras of $C(X)$ containing constant functions. We consider the problem of approximation of a function $f \in C(X)$ by elements from $A_1 + A_2$. We prove a Chebyshev-type alternation theorem for a function $u_0 \in A_1 + A_2$ to be a best approximation to $f$.

Keywords: Chebyshev alternation theorem; best approximation; bolt; weak* convergence; Banach–Alaoglu theorem

2020 Mathematics subject classification: Primary 41A30; 41A50; 46B50; 46E15

1. Introduction

The classical Chebyshev alternation theorem gives a criterion for a polynomial $P$ of degree not greater than $n$ to be the best uniform approximation to a continuous real-valued function $f$, using the oscillating nature of the difference $f - P$. More precisely, the theorem asserts that $P$ is the best uniform approximation to $f$ on $[0, 1]$ if and only if there exist $n + 2$ points $t_i$ in $[0, 1]$ such that

$$f(t_k) - P(t_k) = (-1)^k \max_{t \in [0, 1]} |f(t) - P(t)|, \quad k = 1, \ldots, n + 2.$$ 

See the monograph of Natanson [14] for a comprehensive commentary on this theorem. Several general alternation theorems applying to an arbitrary finite dimensional subspace $M$ of $C(X)$ for $X$, a cell in $\mathbb{R}^d$, may be found in Buck [4]. For the history and various variants of the Chebyshev alternation theorem, consult [3].

In this paper, we prove a Chebyshev-type alternation theorem for a best approximation of a continuous function, defined on a compact metric space, by sums of two algebras. To make the problem more precise, assume $X$ is a compact metric space, $C(X)$ is the space...
of real-valued continuous functions on $X$, $A_1$ and $A_2$ are closed subalgebras of $C(X)$ containing constants. For a given function $f \in C(X)$, consider the approximation of $f$ by elements of $A_1 + A_2$. We ask and answer the following question: which conditions imposed on $u_0 \in A_1 + A_2$ are necessary and sufficient for the equality

$$\|f - u_0\| = \inf_{u \in A_1 + A_2} \|f - u\|? \quad (1.1)$$

Here $\|\cdot\|$ denotes the standard uniform norm in $C(X)$. Recall that a function $u_0$ satisfying Equation (1.1) is called a best approximation to $f$.

It should be remarked that approximation problems concerning sums of algebras were studied in many papers (see Khavinson’s monograph [10] for an extensive discussion). The history of this subject goes back to 1937 and 1948 papers by M.H. Stone [19, 20]. He considered the most particular case of the approximation by sums of algebras, namely the case when only one algebra is involved. A version of the corresponding famous result, known as the Stone–Weierstrass theorem, states that a subalgebra $A \subset C(X)$, which contains a non-zero constant function, is dense in the whole space $C(X)$ if and only if $A$ separates points of $X$ (i.e., for any two different points $x$ and $y$ in $X$, there exists a function $g \in A$ with $g(x) \neq g(y)$). Density of the sum of two subalgebras $A_1$ and $A_2$ in $C(X)$ (for a compact Hausdorff $X$) was extensively studied in Marshall and O’Farrell [12, 13]. In [13], they gave a complete description of measures on $X$ orthogonal to the sum $A_1 + A_2$. From this description, they obtained a geometrical condition, which is equivalent to the density of $A_1 + A_2$ in $C(X)$. The paper [13] also indicates main difficulties with the sum of more than two algebras.

This paper exploits the same mathematical and geometrically explicit objects from Marshall and O’Farrell [12, 13] for characterization of a best approximation by a sum of two algebras. To prove our main result, we use various results and ideas of Functional Analysis and General Topology.

Note that the algebras $A_i$, in particular cases, turn into algebras of univariate functions, ridge functions and radial functions. The literature abounds with the use of ridge functions and radial functions. Ridge functions and radial functions are defined as multivariate functions of the form $g(a \cdot x)$ and $g(|x - a|_e)$, respectively, where $a \in \mathbb{R}^d$ is a fixed vector, $x \in \mathbb{R}^d$ is the variable, $a \cdot x$ is the usual inner product, $|x - a|_e$ is the Euclidean distance between $x$ and $a$ and $g$ is a univariate function.

2. The main result

Let $X$ be a compact metric space, $C(X)$ be the space of real-valued continuous functions on $X$ and $A_1 \subset C(X)$, $A_2 \subset C(X)$ be two closed algebras that contain the constants. Define the equivalence relation $R_i$, $i = 1, 2$, for elements in $X$ by setting

$$a \sim_i b \quad \text{if} \quad f(a) = f(b) \quad \text{for all} \quad f \in A_i.$$

Then, for each $i = 1, 2$, the quotient space $X_i = X/R_i$ with respect to the relation $R_i$, equipped with the quotient space topology, is compact and the natural projections $s : X \to X_1$ and $p : X \to X_2$ are continuous. Note that the quotient spaces $X_1$ and $X_2$ are not only compact but also Hausdorff (see, e.g., [10, p.54]). In view
of the Stone–Weierstrass theorem, the algebras $A_1$ and $A_2$ have the following set representations:

$$A_1 = \{ g(s(x)) : g \in C(X_1) \},$$

$$A_2 = \{ h(p(x)) : h \in C(X_2) \}.$$  

We proceed with the definition of *lightning bolts* with respect to two algebras. These objects are essential for our further analysis.

**Definition 2.1.** (see [13]). A finite or infinite ordered set $l = \{ x_1, x_2, \ldots \} \subset X$, where $x_i \neq x_{i+1}$, with either $s(x_1) = s(x_2), \ p(x_2) = p(x_3), \ s(x_3) = s(x_4), \ldots$ or $p(x_1) = p(x_2), \ s(x_2) = s(x_3), \ p(x_3) = p(x_4), \ldots$ is called a lightning bolt with respect to the algebras $A_1$ and $A_2$.

In the sequel, we will simply use the term “bolt” instead of the expression “lightning bolt with respect to the algebras $A_1$ and $A_2$”. If in a finite bolt $\{ x_1, \ldots, x_n, x_{n+1} \}$, $x_{n+1} = x_1$ and $n$ is an even number, then the bolt $\{ x_1, \ldots, x_n \}$ is said to be closed.

Bolts, in the special case when $X \subset \mathbb{R}^2$, and $A_1 = \{ g(x) \}$ and $A_2 = \{ h(y) \}$ are geometrically explicit objects. In this case, a bolt is an ordered set $\{ x_1, x_2, \ldots \}$ in $\mathbb{R}^2$ with the line segments $[x_i, x_{i+1}], i = 1, \ldots, n$, perpendicular alternatively to the $x$ and $y$ axes. Bolts, in this particular and simplest case, were first introduced by Diliberto and Straus in [6]. They were further used in many works devoted to the approximation of multivariate functions by sums of univariate functions (see [10]). Bolts appeared in a number of papers with several different names such as permissible lines (see [6]), paths (see, e.g., [11]), trips (see, e.g., [12]) and links (see, e.g., [5]). The term bolt of lightning is due to Arnold [1]. Marshall and O’Farrell [13] generalized these objects to the case of two abstract subalgebras of the space of continuous functions defined on a compact Hausdorff space. They gave many central properties of bolts and functionals associated with them.

Let us now define *extremal bolts*.

**Definition 2.2.** A finite or infinite bolt $\{ x_1, x_2, \ldots \}$ is said to be extremal for a function $f \in C(X)$ if $f(x_i) = (-1)^i \| f \|, \ i = 1, 2, \ldots$, or $f(x_i) = (-1)^{i+1} \| f \|, \ i = 1, 2, \ldots$.

We continue with the notion of image of a finite signed measure $\mu$ and a measure space $(U, A, \mu)$. Let $F$ be a mapping from the set $U$ to the set $T$. Then a measure space $(T, B, \nu)$ is called an image of the measure space $(U, A, \mu)$ if the measurable sets $B \in B$ are the subsets of $T$ such that $F^{-1}(B) \in A$ and

$$\nu(B) = \mu(F^{-1}(B)), \quad \text{for all } B \in B.$$  

The measure $\nu$ is called an image of $\mu$ and denoted by $F \circ \mu$. Clearly,

$$\| F \circ \mu \| \leq \| \mu \|,$$  

since under mapping $F$, there is a possibility of mixing up the images of those sets on which $\mu$ is positive with those where it is negative. Besides, note that if a bounded
A measure $\mu \in C^*(X)$ is orthogonal to $A_1 + A_2$ if and only if $s \circ \mu \equiv 0$ and $p \circ \mu \equiv 0$.

That is, for any Borel subsets $E_i \subset X_i$, $i = 1, 2$, $\mu(s^{-1}(E_1)) = 0$ and $\mu(p^{-1}(E_2)) = 0$.

The proof of this lemma easily follows from Equation (2.1).

Lemma 2.2. The quotient spaces $X_1$ and $X_2$ are metrizable.

This lemma is a consequence of the following two facts:

(1) Let $A$ be a family of functions continuous on a compact space $X$ and $r$ an equivalence relation defined by $A$: $x \sim r y$ if $f(x) = f(y)$ for all $f \in A$.

Then the saturation $r(F)$ of any closed set $F \subset X$ ($r(F) \overset{df}{=} \bigcup_{x \in F} r(x)$; $r(x)$ is the equivalence class of $x$) is closed, and hence the canonical projection $\pi : X \to X/r$ is a closed mapping (see [10, p. 54]).

(2) Let $\pi$ be a closed continuous mapping of a metric space $X$ onto a topological space $Y$. Then the following statements are all equivalent (see [18] and [7, Theorem 5.5]):

(a) $Y$ satisfies the first countability axiom.
(b) $\text{bd}(\pi^{-1}(y))$ is compact for each $y \in Y$ (here $\text{bd}(\pi^{-1}(y))$ denotes the boundary of $\pi^{-1}(y)$).
(c) $Y$ is metrizable.

The following theorem plays an essential role in the proof of our main result.

Theorem 2.1. (see Singer [16]). Let $X$ be a compact space, $M$ be a linear subspace of $C(X)$, $f \in C(X) \setminus M$ and $u_0 \in M$. Then $u_0$ is a best approximation in $M$ to $f$ if and only if there exists a regular Borel measure $\mu$ on $X$ such that

1. The total variation $\|\mu\| = 1$;
2. $\mu$ is orthogonal to the subspace $M$, that is, $\int_X u \, d\mu = 0$ for all $u \in M$;
3. For the Jordan decomposition $\mu = \mu^+ - \mu^-$,

$$f(x) - u_0(x) = \begin{cases} \|f - u_0\| & \text{for } x \in \text{supp}(\mu^+), \\ -\|f - u_0\| & \text{for } x \in \text{supp}(\mu^-). \end{cases}$$
where $\text{supp}(\mu^+)$ and $\text{supp}(\mu^-)$ are closed supports of the positive measures $\mu^+$ and $\mu^-$, respectively.

Our main result is the following theorem.

**Theorem 2.2.** Assume $X$ is a compact metric space. A function $u_0 \in A_1 + A_2$ is a best approximation to a function $f \in C(X)$ if and only if there exists a closed or infinite bolt extremal for the function $f - u_0$.

**Proof.** Necessity. Assume $u_0$ is a best approximation from $A_1 + A_2$ to $f$. Since $A_1 + A_2$ is a subspace of $C(X)$, we have a regular Borel measure $\mu$ satisfying the conditions (1)–(3) of Theorem 2.1.

Take any point $x_0$ in $\text{supp}(\mu^+)$ and consider the point $y_0 = s(x_0)$ in $X_1$. Since by Lemma 2.2, $X_1$ is metrizable (hence first countable), there is a nested countable open neighbourhood basis at $y_0$. Denote this basis by $\{O_n(y_0)\}_{n=1}^{\infty}$. For each $n$, $\mu^+[s^{-1}(O_n(y_0))] > 0$, since $s^{-1}(O_n(y_0))$ is an open set containing $x_0$. By Lemma 2.1, $\mu[s^{-1}(O_n(y_0))] = 0$. Therefore, $\mu^-[s^{-1}(O_n(y_0))] > 0$. It follows that for each $n$, the intersection $s^{-1}(O_n(y_0)) \cap \text{supp}(\mu^-)$ is not empty. Take now any points $z_n \in s^{-1}(O_n(y_0)) \cap \text{supp}(\mu^-)$, $n = 1, 2, \ldots$. Since $\text{supp}(\mu^-)$ is sequentially compact (as a closed set in a compact metric space), the sequence $\{z_n\}_{n=1}^{\infty}$ or a subsequence of it converges to a point $x_1$ in $\text{supp}(\mu^-)$. We may assume without loss of generality that $z_n \to x_1$, as $n \to \infty$. Since for any $n$, $s(z_n) \in O_n(y_0)$ and $\{O_n(y_0)\}_{n=1}^{\infty}$ is a nested neighbourhood basis, we obtain that $s(z_n) \to y_0$, as $n \to \infty$. On the other hand, since $s$ is continuous, $s(z_n) \to s(x_1)$, as $n \to \infty$. It follows that $s(x_1) = y_0 = s(x_0)$. Note that $x_0 \in \text{supp}(\mu^+)$ and $x_1 \in \text{supp}(\mu^-)$.

Changing $s$ and $\mu^+$ to $p$ and $\mu^-$, correspondingly, repeat the above process with the point $y_1 = p(x_1)$ and a nested countable neighbourhood basis at $y_1$. Then we obtain a point $x_2 \in \text{supp}(\mu^+)$ such that $p(x_2) = p(x_1)$. Continuing this process, we can construct points $x_3 \in \text{supp}(\mu^-)$, $x_4 \in \text{supp}(\mu^+)$, and so on. Note that the set of all constructed points $x_i$, $i = 0, 1, \ldots$, forms a bolt. By Theorem 2.1, this bolt is extremal for the function $f - u_0$.

Sufficiency. The main idea in this part is the application of the Banach–Alaoglu theorem on weak* sequential compactness of the closed unit ball in $E^*$ for a separable Banach space $E$ (see, e.g., Rudin [15, p. 66]). Note that since $X$ is a compact metric space, the space $C(X)$ is separable. Thus, the closed unit ball $B$ of the continuous dual of $C(X)$ is sequentially compact, which means that any sequence in $B$ has a converging subsequence converging to a point in $B$.

With each bolt $l = \{x_1, \ldots, x_n\}$ with respect to $A_1$ and $A_2$, we associate the following bolt functional

$$r_l(F) = \frac{1}{n} \sum_{i=1}^{n} (-1)^{n+1} F(x_i).$$

It is an exercise to check that $r_l$ is a linear bounded functional on $C(X)$ with the norm $\|r_l\| = 1$ and $\|r_l\| = 1$ if and only if the set of points $x_i$ with odd indices $i$ does not intersect with the set of points with even indices. Besides, if $l$ is closed, then $r_l \in (A_1 + A_2)^\perp$, where $(A_1 + A_2)^\perp$ is the annihilator of the subspace $A_1 + A_2 \subset C(X)$. If $l$ is not closed, then $r_l$
A. K. Asgarova, A. A. Huseynli and V. E. Ismailov

is generally not an annihilating functional. However, it satisfies the following important inequality

$$|r_l(v_i)| \leq \frac{2}{n} \|v_i\|,$$  \hspace{1cm} (2.2)

for all $v_i \in A_i$, $i = 1, 2$. This inequality means that for bolts $l$ with sufficiently large number of points, $r_l$ behaves like an annihilating functional on each $A_i$, and hence on $A_1 + A_2$. To see the validity of Equation (2.2), it is enough to recall that $v_1 = g \circ s$, $v_2 = h \circ p$ and consider the chain of equalities $g(s(x_1)) = g(s(x_2))$, $g(s(x_3)) = g(s(x_4))$, $\ldots$ (or $g(s(x_2)) = g(s(x_3))$, $g(s(x_4)) = g(s(x_5))$, $\ldots$) for $v_1(x) = g(s(x))$ and similar equalities for $v_2(x) = h(p(x))$.

Returning to the sufficiency part of the theorem, note that there may be two cases. The first case happens when there exists a closed bolt $l = \{x_1, \ldots, x_{2n}\}$ extremal for $f - u_0$. In this case, it is not difficult to verify that $u_0$ is a best approximation. Indeed, on the one hand, the following equalities are valid:

$$|r_l(f)| = |r_l(f - u_0)| = \|f - u_0\|.$$

On the other hand, for any function $u \in A_1 + A_2$, we have

$$|r_l(f)| = |r_l(f - u)| \leq \|f - u\|.$$

Thus, $\|f - u_0\| \leq \|f - u\|$ for any $u \in A_1 + A_2$. That is, $u_0$ is a best approximation.

The second case is the existence of an infinite bolt $l = \{x_1, x_2, \ldots\}$ extremal for $f - u_0$. In this case, we proceed as follows. From $l$, we form the finite bolts $l_k = \{x_1, \ldots, x_k\}$, $k = 1, 2, \ldots$, and consider the bolt functionals $r_{l_k}$. For the ease of notation, let us put $r_k = r_{l_k}$. The sequence $\{r_k\}_{k=1}^{\infty}$ is contained in the closed unit ball of the dual space $C^*(X)$. By the Banach–Alaoglu theorem, the sequence $\{r_k\}_{k=1}^{\infty}$ must have weak* cluster points. Assume $r^*$ is one of them. Without loss of generality, we may assume that $r_k \xrightarrow{\text{weak*}} r^*$, as $k \to \infty$. From Equation (2.2) it follows that $r^*(v_1 + v_2) = 0$, for any $v_i \in A_i$, $i = 1, 2$. That is, $r^*$ belongs to the annihilator of the subspace $A_1 + A_2$. Since we have also $\|r^*\| \leq 1$, it follows that

$$|r^*(f)| = |r^*(f - u)| \leq \|f - u\|$$  \hspace{1cm} (2.3)

for all functions $u \in A_1 + A_2$. On the other hand, since the infinite bolt $\{x_1, x_2, \ldots\}$ is extremal for $f - u_0$,

$$|r_k(f - u_0)| = \|f - u_0\|, k = 1, 2, \ldots.$$

Hence,

$$|r^*(f)| = |r^*(f - u_0)| = \|f - u_0\|.$$  \hspace{1cm} (2.4)
From Equations (2.3) and (2.4), we obtain that

$$\| f - u_0 \| \leq \| f - u \|$$

for all $u \in A_1 + A_2$. This means that $u_0$ is a best approximation to $f$. \hfill $\square$

**Remark 1.** In [2], Theorem 2.2 was proved under additional assumption that the algebras have the $C$-property, that is, for any $w \in C(X)$, the functions

$$g_1(a) = \max_{s(x) = a} w(x), g_2(a) = \min_{s(x) = a} w(x), a \in X_1,$$

$$h_1(b) = \max_{p(x) = b} w(x), h_2(b) = \min_{p(x) = b} w(x), b \in X_2$$

are continuous.

**Remark 2.** Note that in the special case when $X \subset \mathbb{R}^2$ and $s, p$ are the coordinate functions, a Chebyshev-type alternation theorem was first obtained by Havinson [9]. In [8], similar alternation theorems were proved for ridge functions and certain function compositions.

**Remark 3.** Note that characterization of a best approximation from a sum of more than two subalgebras $A_1, \ldots, A_k$ of $C(X)$ seems to be beyond the scope of the methods discussed herein. A bolt with respect to two algebras $A_1$ and $A_2$ is constructed as a sequence of points $\{x_1, x_2, \ldots\}$ with the links $x_i x_{i+1}$ travelling alternatively in equivalence classes of the relations $R_1$ and $R_2$ (see above). In this case, the bolt functional $r_1$ has important property (2.2), which leads to the functional $r^*$ annihilating all elements of the sum $A_1 + A_2$. The problem becomes complicated when the number of summands in the sum $A_1 + \cdots + A_k$ is more than two. The simple generalization of bolts demands a sequence of points $\{x_1, x_2, \ldots\}$ with the links $x_i x_{i+1}$ travelling in three or more alternating equivalence classes. But in this case, the number 2 in Equation (2.2) grows unboundedly as $n$ tends to infinity, and we cannot arrive at any annihilating functional like $r^*$. For $k \geq 3$, we do not know a reasonable description of a sequence of points $\{x_1, x_2, \ldots\}$ and functionals $r_{l_n}$, associated with the first $n$ points $x_1, \ldots, x_n$, such that any weak$^*$ cluster point of the sequence $\{r_{l_n}\}_{n=1}^{\infty}$ is orthogonal to the sum $A_1 + \cdots + A_k$. We refer the interested reader to Sternfeld [17] for discussions on differences between the cases of two and more than two algebras.

**Competing interests.** The authors declare none.

**References**

(1) V. I. Arnold, On functions of three variables, (Russian), *Dokl. Akad. Nauk SSSR* 114(2) (1957), 679–681. English transl. in Amer. Math. Soc. Transl. 28 (1963), 51–54.

(2) A. K. Asgarova and V. E. Ismailov, A Chebyshev-type theorem characterizing best approximation of a continuous function by elements of the sum of two algebras, (Russian, *Mat. Zametki* 109(1) (2021), 19–26).
(3) B. Brosowski and A. R. da Silva, *A general alternation theorem, approximation theory* (Memphis, TN, 1991), 137–150. Lecture Notes in Pure and Applied Mathematics, Volume 138 (Dekker, New York, 1992).

(4) R. C. Buck, Alternation theorems for functions of several variables, *J. Approx. Theory* 1(3) (1968), 325–334.

(5) R. C. Cowsik, A. Klopotowski and M. G. Nadkarni, When is \( f(x, y) = u(x) + v(y) \)?, *Proc. Indian Acad. Sci. Math. Sci.* 109(1) (1999), 57–64.

(6) S. P. Diliberto and E. G. Straus, On the approximation of a function of several variables by the sum of functions of fewer variables, *Pacific J. Math.* 1(2) (1951), 195–210.

(7) R. A. Herman, *Quotients of metric spaces* (A.B. Grinnell College, 1966).

(8) V. E. Ismailov, *Ridge functions and applications in neural networks*. Mathematical Surveys and Monographs 263, p. 186, (American Mathematical Society, 2021).

(9) S. J. Havinson, A Chebyshev theorem for the approximation of a function of two variables by sums of the type \( \varphi(x) + \psi(y) \), *Izv. Acad. Nauk. SSSR Ser. Mat.* 33(3) (1969), 650–666. English transl. in *Math. USSR Izv.* 3 (1969), 617–632.

(10) S. Y. Khavinson, Best approximation by linear superpositions (approximate nomography), Translated from the Russian manuscript, D. Khavinson (ed) *Translations of mathematical monographs*, Volume 159, p. 175 (American Mathematical Society, Providence, RI, 1997).

(11) W. A. Light and E. W. Cheney, On the approximation of a bivariate function by the sum of univariate functions, *J. Approx. Theory* 29(4) (1980), 305–322.

(12) D. E. Marshall and A. G. O’Farrell, Uniform approximation by real functions, *Fund. Math.* 104(3) (1979), 203–211.

(13) D. E. Marshall and A. G. O’Farrell, Approximation by a sum of two algebras. The lightning bolt principle, *J. Funct. Anal.* 52(3) (1983), 353–368.

(14) I. P. Natanson, Constructive function theory, Vol. I. Uniform approximation, A. N. Obolensky (ed), *Translated from the Russian*, p. 232 (Frederick Ungar Publishing Co., New York, 1964).

(15) W. Rudin, *Functional analysis, McGraw-Hill Series in Higher Mathematics*, p. 397 (McGraw-Hill Book Co, 1973).

(16) I. Singer, *The theory of best approximation and functional analysis*, Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 13. Society for Industrial and Applied Mathematics, Philadelphia, PA, p. 95 (SIAM Publications Library, 1974).

(17) Y. Sternfeld, Uniform separation of points and measures and representation by sums of algebras, *Israel J. Math.* 55(3) (1986), 350–362.

(18) A. H. Stone, Metrizability of decomposition spaces, *Proc. Amer. Math. Soc.* 7(4) (1956), 690–700.

(19) M. H. Stone, Applications of the theory of Boolean rings to general topology, *Trans. Amer. Math. Soc.* 41(3) (1937), 375–481.

(20) M. H. Stone, The generalized Weierstrass approximation theorem, *Math. Mag.* 21(4) (1948), 167–184, 237–254.