James H. McClellan
and Thomas W. Parks

A Personal History of the Parks–McClellan Algorithm

The subject of finite impulse response (FIR) filter design was one of the hottest research topics in DSP in the late 1960s and early 1970s. Several groups were, in effect, competing to write the best design program that would produce optimal filters for any possible specification and for any conceivable filter length. This article describes the work that led to what is now known as the Parks-McClellan algorithm [1]. Within the bigger picture of filter design methods, our work was just one story, so we will only recount events that had an impact on our inspiration to develop the Parks-McClellan algorithm, i.e., the Remez exchange algorithm with optimal Chebyshev approximation for FIR filter design. Furthermore, these recollections are those that have survived for the past 30 years as we look back on what was, for both of us, one of the most stimulating research quests of our careers (even though we didn't realize it at the time).

Before describing the details of our thinking as we worked on the algorithm for optimal filter design, it is worthwhile to lay out a timeline for the events that led to the development of the Parks-McClellan algorithm and our paper [1]. It is astounding how quickly the events unfolded during the spring of 1971. In March of that year, Tom Parks attended a conference in Princeton, New Jersey, and heard Ed Hofstetter's presentation about a new FIR filter design algorithm [2]. During Tom's spring semester course, Hofstetter's paper became a programming project for Jim McClellan. At the end of the semester, the quest was on to write a variation of the Remez exchange algorithm for FIR filters. This took about six weeks (for reasons that will be disclosed later); by the end of May, some optimal filters had been designed successfully. (It should be mentioned that on 25 May, Amy McClellan was born, so May of 1971 was indeed memorable).

With a working program, both of us had the feeling of a true research triumph. However, since it was likely that several other researchers were very close to the same answer, we felt that a rapid publication was needed to assert the first claim. In those days, *Electronics Letters* provided quick turn-around on articles, but the article had to be in the form of a letter limited to about five type-written pages. We rapidly dashed off that letter and set about writing a full paper. Imagine our surprise when a couple of weeks later we received a rejection from *Electronics Letters*. The reviewers didn't believe our claims, but we knew that we had the filters to prove we were right. We continued with our full paper and submitted it on 29 June 1971 to *IEEE Transactions on Circuit Theory*. The first, and most important, chapter of the Parks-McClellan algorithm was complete, and its development had taken about three months.

The full story began in August 1970, when Jim McClellan entered graduate school at Rice thinking that the mathematical methods of analog filter design or network synthesis might be an interesting area for study. However, two young professors had just launched an exciting new course called "Digital Filters" the year before. Since it was the closest thing to analog filters, McClellan enrolled. The textbook was coauthored by Ben Gold and Charlie Rader, and the course was taught jointly by Tom Parks and Sid Burrus. At that time, DSP was such a young field that lectures sometimes involved recently published research papers and homework problems sometimes became new journal papers to write.

The following semester (spring 1971), Tom Parks offered a course called "Signal Theory," which covered function spaces and approximation theory (representing signals with basis functions). During spring break, Tom drove from Houston to Princeton to attend the conference mentioned earlier. He brought back the paper by Hofstetter, Oppenheim, and Siegel [2], excited about the possibility of using Chebyshev approximation theory to design FIR filters. The students in the "Signal Theory" course were required to do a project, and since Chebyshev approximation was a major topic in the course, the implementation of this new algorithm became Jim McClellan's course project.

If we go back in time one or two decades prior to 1971, it is possible to find many instances where Chebyshev approximation was used for filter design. Practitioners of analog filter design in the 1960s were well aware that the best filters exhibited an equiripple characteristic in their frequency response magnitude, so designers used elliptic (or Cauer) filters when tight specifications on the magnitude response were required. The goal was to minimize the "worst-case" error in the passband and stopbands, which led to the mathematical theory of Chebyshev (or min–max) approximation. Even though the theory is not simple for rational transfer functions, it was generally acknowledged that the elliptic filter was *optimal in the Chebyshev sense*.

When the digital filter revolution began in the 1960s, researchers used the

1053-5888/05/$20.00©2005IEEE

bilinear transformation to produce infinite impulse response (IIR) digital elliptic filters; they also recognized the potential for designing FIR filters to accomplish the same filtering task. Soon the search was on for the optimal FIR filter *in the Chebyshev sense*. The theory for the FIR design problem is more powerful than for the IIR case because the FIR problem is linear in the unknown filter coefficients. Therefore, it was well known in both mathematics and engineering that the optimal response would exhibit an equiripple behavior and, furthermore, that the number of ripples could be counted (based on Chebyshev's work done in the 1800s). Simply stated, the number of ripples has to be one greater than the degrees of freedom in the problem.

Several attempts to produce a design program for optimal Chebyshev FIR filters were undertaken in the period 1962–1971. Most of these had some small defect in either the problem formulation or the algorithmic implementation, which left them unable to produce a general optimal filter design method. Otto Herrmann, working on his thesis under the direction of Hans Schüssler, proposed a method for designing equiripple filters with restricted band edges [3]. Their method obtained an equiripple frequency response with the maximum number of ripples by solving a set of nonlinear equations. Another method that was based on linear programming did provide an optimal Chebyshev approximation, but it led to an algorithm that was limited to the design of relatively low-order filters unless a very powerful computer was used.

Similar to the Herrmann and Schüssler's method, the algorithm presented by Hofstetter [2] (which became Jim McClellan's course project) designed FIR filters with as many ripples as possible. Here it will be referred to as the Maximal Ripple (MR) algorithm. The MR algorithm (see "Maximal Ripple Algorithm") imposed an alternating error condition via interpolation and then solved a set of equations that the alternating solution had to satisfy. One notable limitation of the MR algorithm was that the band edges (the end of the passband and beginning of the stopband) were not specified as inputs to the design procedure. For instance, it was not possible to place the passband edge exactly at $\omega = 0.2\pi$ and the stopband edge exactly at $\omega = 0.3\pi$. Rather,

In this issue, we celebrate the "DSP History" column's first anniversary. In March 2004, when Larry Rabiner's "The CZT Transform: A Lesson in Serendipity" marked the joyous birth of the column, we embarked on a journey that proved to be far more stimulating, challenging, and rewarding than we could have ever predicted. During this journey, we have had special guests who brought wisdom, inspiration, and humor to the column, and helped us all understand major DSP events in a broader professional and personal context. We thank them all and our readers alike. What better way to celebrate could there be than to continue to rediscover and understand the details associated with historical DSP events. This issue's column offers insights into the well-known Parks-McClellan algorithm and the over three-decade collaboration of its authors.

James McClellan was born on 5 October 1947 in Guam. He obtained his B.S.E.E. (1969) from Louisiana State University and his M.Sc. (1972) and Ph.D. (1973) degrees from Rice University. He has been with the MIT Lincoln Laboratory (1973–1975), MIT Electrical Engineering and Computer Science Department (1975–1982), and Schlumberger (1982–1987). Since 1987, he has been a professor of electrical engineering at the Georgia Institute of Technology. His work has spanned various aspects of digital signal processing with application to sensor array processing (radar, geophysics, and acoustic sensing). Dr. McClellan coauthored *Computer–Based Exercises for Signal Processing Using MATLAB 5* (1994), *DSP First* (1997), *Signal Processing First* (2003), and *Number Theory in DSP* (1979). He is the recipient of the IEEE Signal Processing Technical Achievement Award (1987), the IEEE Signal Processing Society Award (1996), the IEEE Jack S. Kilby Signal Processing Medal (2004), and several teaching awards. His hobbies include running and reading books about history and sports. With a strong focus on the positive aspects of life, he appreciates lively discussions about interesting ideas, especially if shared along with good wine and good food.

Thomas Parks was born on 16 March 1939 in Buffalo, New York. He received his B.E.E. (1961), M.Sc. (1964), and Ph.D. (1967) degrees in electrical engineering, all from Cornell University. From 1967 to 1986, he was a faculty member at Rice University in Houston. Since 1986, he has been a professor of electrical engineering with the School of Electrical and Computer Engineering at Cornell University. His work has focused on digital signal processing with application to signal theory, multirate systems, interpolation, and filter design. Dr. Parks coauthored *DFT/FFT and Convolution Algorithms* (1985), *Digital Filter Design* (1987), *A Digital Signal Processing Laboratory Using the TMS 32010* (1988), *A Digital Signal Processing Laboratory Using the TMS 320C25* (1990), *Computer–Based Exercises for Signal Processing* (1994), and *Computer–Based Exercises for Signal Processing Using MATLAB 5* (1994). He is the recipient of the IEEE ASSP Society Technical Achievement Award (1981), the IEEE ASSP Society Award (1988), the Rice University President's Award (1999), the IEEE Third Millennium Medal (2000), the IEEE Jack S. Kilby Signal Processing Medal (2004), and numerous teaching awards. His happiest professional moment was receiving the Kilby Medal with Jim McClellan in June 2004. "My idea man" (as Tom Parks was nicknamed by Jim McClellan) enjoys golf, sailing, and flying radio control model airplanes, preferably all in sunshine.

In the personal history told here, Jim McClellan and Tom Parks show us that time has been gentle: over the three decades that have passed since their discovery of the Parks-McClellan algorithm, the details and people surrounding it have gained permanent places in personal and collective memories.

—Adriana Dumitras, George Moschytz
"DSP History" column editors
adrianad@ieee.org
moschytz@isi.ee.ethz.ch

MAXIMAL RIPPLE ALGORITHM

The Maximal Rripple (MR) algorithm formulated the filter design problem by requiring as many ripples as possible. It also imposed an alternating error via an interpolation condition. Instead of directly minimizing the worst–case error, the MR algorithm solved a set of equations that the alternating solution had to satisfy

$$D(\omega_i) + (-1)^i \epsilon = \sum_{k=0}^{M} a_k \cos(\omega_i k) \tag{1}$$

where $D(\omega_i)$ and $|\epsilon|$ were the desired frequency response of the filter and the desired Chebyshev error, respectively. The term $(-1)^i$ forced the error to alternate on the set of frequencies $\{\omega_i\}$. This set $\{\omega_i\}$ included both ends of the frequency axis, $\omega = 0$ and $\omega = \pi$. (1) contains $M + 1$ unknown coefficients $\{a_k\}$. Once $M + 1$ frequencies are chosen in (1), the coefficients $\{a_k\}$ can be found by solving a set of simultaneous linear equations.

Then the question is how to find the frequencies $\{\omega_i\}$. With knowledge of the coefficients $\{a_k\}$, the error $E(\omega)$ over the entire frequency domain can be evaluated as the difference between the desired function $D(\omega)$ and its approximation

$$E(\omega) = D(\omega) - \sum_{k=0}^{M} a_k \cos(\omega k). \tag{2}$$

The filter length is $L = 2M + 1$, which must be odd for this form to hold. If there are points where $|E(\omega)| > \epsilon$, (i.e. , locations where the actual maximum error exceeds the desired maximum error), then the frequencies $\{\omega_i\}$ have to be adjusted. The MR strategy was to form a new set of frequencies using the locations where $E(\omega)$ reached its local maxima, and then iterate until $|E(\omega)| \le \epsilon$.

ALTERNATION THEOREM

The Alternation Theorem is most often stated for approximating a continuous function over an interval. However, in [5] Cheney gave a version that is true for any closed subset of an interval $\Omega \subset [a, b]$. The function being approximated must be *continuous on* Ω; after a bit of thought, it is easy to see that a desired function such as

$$D(\omega) = \begin{cases} 1 & 0 \le \omega \le \omega_p \\ 0 & \omega_s \le \omega \le \pi \end{cases} \tag{3}$$

is, in fact, a continuous function on the closed set $\Omega = [0, \omega_p] \cup [\omega_s, \pi]$, where ω_p and ω_s are the band edges (the end of the passband and the beginning of the stopband, respectively). With a continuous function $D(\omega)$, the theory of Chebyshev approximation is quite clear on the necessary and sufficient conditions for having an optimal filter. The famous *Alternation Theorem of Chebyshev approximation* states that the error function $E(\omega)$ in (2) will reach its maximum absolute value *at least* $M + 2$ times, (ie., one more than the number of cosine functions in the FIR frequency response). In addition, if the frequencies at which the error reaches its maximum absolute value are ordered

$$\omega_1 < \omega_2 < \cdots < \omega_{M+2} \tag{4}$$

then the maximum error alternates, meaning that

$$|E(\omega_{i+1})| = -|E(\omega_i)| \quad \text{for } i = 1, 2, \ldots, M + 1. \tag{5}$$

The MR algorithm [2] and the Herrmann-Schüssler method [3] created filters for which the error had equal ripples, at frequencies where the error achieved its maximum magnitude. In fact, the count was the $M + 1$ frequencies inside the frequency band of $[0, \pi]$, plus the two frequencies, $\omega = 0$ and $\omega = \pi$, giving a total of $M + 3$ frequencies—one more than required by the Alternation Theorem.

the initial frequency set $\{\omega_i\}$ and the desired function $D(\omega_i)$ defined the passband and stopband implicitly. For example, to design a low-pass filter of length 19 (i.e., $M = 9$), it was necessary to distribute ten frequencies between the two bands. If, in the initial guess of the set $\{\omega_i\}$, the first three frequencies were assigned to the passband and the rest to the stopband, it was reasonable to expect that the transition region might end up near 0.3π; yet, the exact band edges could not be known until the design algorithm completed its iteration. Thus, it did not seem possible to call the MR an optimal filter design.

On the other hand, the MR algorithm used a different type of method than its predecessors to compute the best filter—an exchange algorithm that tried to find the frequency set $\{\omega_i\}$ where the best filter had its ripples. In [2], the authors wrote that their method was "reminiscent of, but different from the Remes exchange algorithm." This led Tom Parks to ask the question "Why not use the Remez exchange algorithm?" The answer required a restatement of the FIR design problem in the precise language of optimal Chebyshev approximation. This ultimately led to the Parks-McClellan algorithm, which involved not only the general theory of optimal Chebyshev approximation, but also an efficient implementation.

Two steps were necessary to make the Chebyshev theory applicable for filter design: first, define the set of basis functions for the approximation, and second, exploit the fact that the passbands and stopbands of bandpass filters would always be separated by *transition regions*. At the beginning, it was assumed that the filter coefficients had even symmetry and the filter length was odd, which led to cosine basis functions. Cosines satisfied the Haar condition, which was needed for the Alternation Theorem (see "Alternation Theorem"). Later work showed that all linear-phase FIR filters could be reduced to the sum of cosines case, so the same core program could be used to perform all possible linear-phase FIR filters. The need for transition regions was well known for

PARKS-MCCLELLAN ALGORITHM

The Parks-McClellan (Remez exchange) algorithm consists of the following steps:

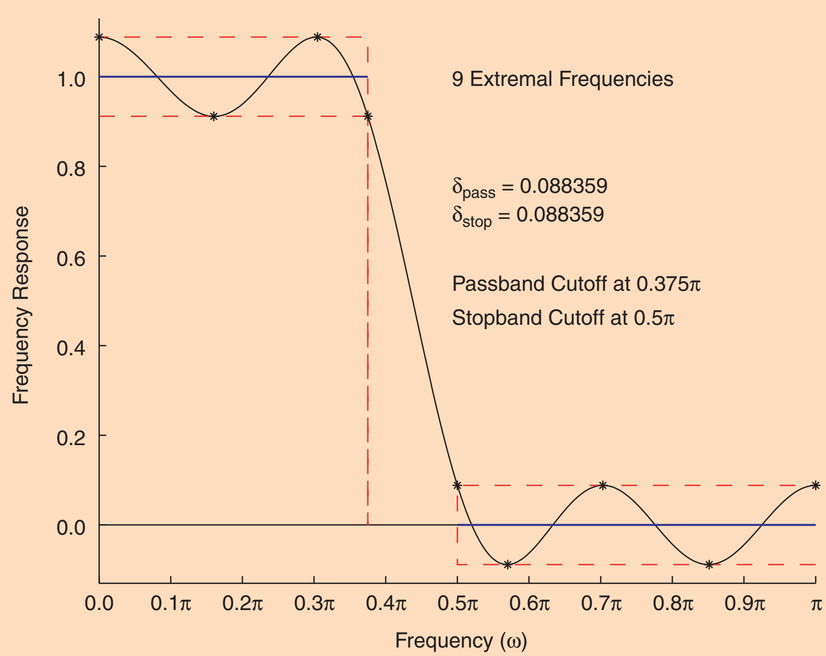
1) Initialization: Choose an extremal set of frequencies $\{\omega_i^{(0)}\}$.
2) Finite Set Approximation: Calculate the best Chebyshev approximation on the present extremal set, giving a value $\delta^{(m)}$ for the min-max error on the present extremal set.
3) Interpolation: Calculate the error function $E(\omega)$ over the entire set of frequencies Ω using (2).
4) Look for local maxima of $|E^{(m)}(\omega)|$ on the set Ω.
5) If $\max_{\omega\in\Omega} |E^{(m)}(\omega)| > \delta^{(m)}$, then update the extremal set to $\{\omega_i^{(m+1)}\}$ by picking new frequencies where $|E^{(m)}(\omega)|$ has its local maxima. Make sure that the error *alternates* on the ordered set of frequencies as described in (4) and (5). Return to Step 2 and iterate.
6) If $\max_{\omega\in\Omega} |E^{(m)}(\omega)| \le \delta^{(m)}$, then the algorithm is complete. Use the set $\{\omega_i^{(m)}\}$ and the interpolation formula to compute an inverse discrete Fourier transform to obtain the filter coefficients.

elliptic filters and the like because the height of the alternating ripples would increase if the transition width were squeezed down. Furthermore, a transition region could be a *don't care* region where the approximation error $E(\omega)$ would be ignored when evaluating the worst-case error. In other words, the maximum of $|E(\omega)|$ would be evaluated on a union of disjoint closed sets consisting of the passbands and stopbands of the desired filter. In contrast to the MR approach, the band edges could now be specified ahead of time.

Once the mathematical statement of the Chebyshev approximation problem was formalized, it was possible to understand how previous work was related to the optimal Chebyshev approximation. If we restrict attention to the low-pass case, the key observation turns out to be that the band edges would always be frequencies where the error reached its maximum absolute value. Once this fact was discovered, the MR design could be reinterpreted as a filter that had a total of $M+3$ alternations, one greater than required by the theorem. These filters were later renamed *extraripple* filters for the low-pass case, or *maximal ripple*, in general.

To achieve an efficient implementation of the optimal filter design using the Parks-McClellan (Remez exchange) algorithm (see "Parks-McClellan Algorithm" and Figure 1), two difficulties had to be overcome: 1) defining a flexible exchange strategy and 2) implementing a robust interpolation method. Several books contained descriptions of the Remez exchange algorithm, but, as expected, these descriptions assumed that the approximation was being conducted over an interval $[a, b]$, whereas we wanted to work over the union of closed subsets of the interval $[0, \pi]$. As luck would have it, the Rice library contained a translation [4] of Remez's original work. This document described several variations of the exchange strategy for doing Chebyshev approximation, and it also confirmed that the method would work for any closed subset of an interval including discrete points. In the computer program, this meant that the passbands and stopbands would be replaced with dense discretized grids covering the same regions.

In some sense, the programming involved the implementation and adaptation of a known algorithm for use in FIR filter design. Two facets of the exchange strategy were tackled to make the program more efficient: 1) allocating the extremal frequencies



[FIG1] Optimal length-15 FIR linear-phase filter designed using the Parks-McClellan algorithm, demonstrating the equiripple nature of the frequency response. The frequency response of the length-15 linear-phase FIR filter can be written as a weighted sum of eight cosine functions $\{1, \cos\omega, \ldots, \cos 7\omega\}$ (i.e., $M = 7$). The Alternation Theorem requires $M+2 = 9$ extremal points (denoted by dots on the frequency response). Both band edges are extremal points, as are $\omega = 0$ and $\omega = \pi$. The optimal frequency response ripples above and below the desired response (in blue) which equals one in the passband $[0, 0.375\pi]$ and zero in the stopband $[0.5\pi, \pi]$.

between the passband and stopband (for the low-pass case) and 2) enabling movement of the extremals between the bands as the program iterated. At initialization, the number of extremals in the passband and stopband could be assigned by using the ratio of the sizes of the bands. Furthermore, the passband edge and stopband edge would always be placed in the extremal set, and the program's logic kept those edge frequencies in the extremal set. The movement between bands was controlled by comparing the size of the errors at all the candidate extremal frequencies and taking the largest.

The second element of the algorithm was the interpolation step needed to evaluate the error function. We used a method called the Barycentric form of Lagrange interpolation (also used by the MR program), which was very robust. The interpolation of very high-order cosine polynomials is a rather sensitive numerical problem, especially when the frequencies are close together. In the filter design problem, the extremal frequencies tend to bunch up near the band edges, so this numerical sensitivity almost always occurs for long filters. Once the program started to be employed for extremely long filters, it was somewhat surprising that the method continued to work. Much of that success can be attributed to the Barycentric interpolation formula.

The actual software development and testing was slow and tedious. It's hard to convey the difference between contemporary software development and programming in the 1970s. Today's interactive debugging environments would have been a fantasy to a graduate student using the batch systems of the early 1970s, when the compile-run-debug cycle involved one run per day and hours of thinking in between to diagnose a program failure. Furthermore, most graduate students of that era had virtually no prior programming experience. In fact, the FIR filter design program was probably McClellan's third or fourth FORTRAN program, and it took about six weeks to debug.

Once we started to distribute the program, it is fair to say that the Parks–McClellan algorithm gained much of its popularity because it was a robust algorithm that worked for many different cases. Much of the credit for the algorithm's robustness goes to Larry Rabiner. Larry was the author of two competing filter design techniques: frequency sampling and a linear programming implementation of Chebyshev approximation. But he jumped on the Remez bandwagon and became a strong advocate for our method once he was convinced that it would live up to its billing.

Nowadays, it would be trivial to think about running an algorithm thousands, or even millions, of times to test its capabilities. But if we once again travel back to the early 1970s, computers were relatively slow and graduate students got only a few minutes of CPU time per day. Exhaustive testing was out of the question. On the other hand, program efficiency was paramount because it was the only way to squeeze out a few more runs on such a small daily budget. The Remez method was very attractive because it was supposed to have quadratic convergence, and we expected the running time to increase as $\mathcal{O}(L^2)$. In contrast, linear programming was known to be $\mathcal{O}(L^3)$, thus limiting itself to filters with lengths less than 100 on the most powerful machines of that time. It is interesting to point out that the first publication of the Remez algorithm for FIR filter design contained a figure that plotted the running time and predicted a trend of $\mathcal{O}(L^{1.626})$ from the empirical timing data.

In the 1970s, Bell Labs had extensive computer resources, and Larry Rabiner was able to run hundreds of filters per day. Once he obtained our program, he started to put it through its paces. One of his objectives was to develop the design rules for the linear phase filters by running thousands of filters and then fitting the results with a relatively simple formula. This led to the approximation formulas that now enable designers to predict the filter length given desired passband and stopband specs on the ripples and band edges.

Initially, Larry's testing led to regular phone calls back to Rice, during which he would report some bug or deficiency that would require corrections. Since the computational power at Bell Labs seemed to be infinite, the questions soon became something like "Why did the program fail for one length-501 filter but work for another?" 501! It was hard to imagine that the interpolation formula would hold up for any length-501 filter, but such questions led to improvements in the numerical robustness of the Barycentric formula by reordering its computations.

From our perspective at Rice, it seemed that Larry wanted to set records for the longest optimal filter ever designed. One day we received a printout of the coefficients of a length-1401 filter; this probably would have consumed several days of CPU time on our batch machine at Rice. Later refinements in the Remez algorithm have addressed the ultra-long filter case, but the primary issue is getting a reasonable initial guess of the extremal frequencies so that the algorithm will get off the ground. Once the value of error on the extremal set rises above the roundoff noise of the computer, the algorithm usually converges.

In conclusion, it's satisfying to recall that the research that led to the Parks-McClellan algorithm involved frequent insights and discoveries that seemed to come almost daily. It was the best kind of research experience because the problem was "hot" and there was an optimal theory that could be applied. We were lucky that so many pieces of the puzzle fell into place so quickly.

REFERENCES

[1] T.W. Parks and J.H. McClellan, "Chebyshev approximation for nonrecursive digital filters with linear phase," *IEEE Trans. Circuit Theory*, vol. CT-19, no. 2, pp. 189–194, 1972.

[2] E.M. Hofstetter, A.V. Oppenheim, and J. Siegel, "A new technique for the design of non-recursive digital filters," in *Proc. 5th Annu. Princeton Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 1971, pp. 64–72.

[3] O. Herrmann, "Design of nonrecursive digital filters with linear phase," *Electron. Lett.*, vol. 6, no. 11, pp. 328–329, 28 May 1970.

[4] E. Ya. Remez, "General computational methods of Tchebycheff approximation," Kiev (Atomic Energy Commission Translation 4491), pp. 1–85, 1957.

[5] E.W. Cheney, *Introduction to Approximation Theory*. New York: McGraw-Hill, 1966, pp. 72–100.

SP