24. Rational best approximation

Chapter 10 considered best or "minimax" approximation by polynomials, that is, approximation in the ∞-norm, where optimality is characterized by an equioscillating error curve. This chapter presents analogous results for approximation by rational functions. Much remains the same, but a crucial new feature is the appearance in the equioscillation condition of a number known as the *defect*, which leads to the phenomenon of square blocks of degenerate entries in the "Walsh table" of best approximations. This complication adds a fascinating new ingredient to the theory, but it is a complication with destructive consequences in terms of the fragility of rational approximations and the difficulty of computing them numerically. An antidote to some such difficulties may be the use of algorithms based on linearized least-squares and the singular value decomposition, a theme we shall take up in Chapters 26 and 27.

Another new feature in rational approximation is that we must now be careful to distinguish real and complex situations, because of a curious phenomenon: best rational approximations to real functions are in general complex. This effect is intriguing, but it has little relevance to practical problems, so for the most part we shall restrict our attention to approximations in the space $\mathcal{R}_{mn}^{\text{real}}$ consisting of functions in $\mathcal{R}_{mn}$ with real coefficients.

We will first state the main theorem, then give some examples, and then present a proof. To begin the discussion, we must define the defect. Suppose $r \in \mathcal{R}_{mn}$, that is, r is a rational function of type (m, n). As discussed in the last chapter, this means that r can be written as a fraction p/q in lowest terms with p and q having exact degrees $\mu \leq m$ and $\nu \leq n$. The *defect* d of r in $\mathcal{R}_{mn}$ is the number between 0 and n defined by

$$d = \min\{m - \mu, n - \nu\} \geq 0. \tag{24.1}$$

Note that d is a measure of how far *both* the numerator and the denominator degrees fall short of their maximum allowed values. Thus $(1 - x^2)/(1 + x^2)$, for example, has defect 0 in $\mathcal{R}_{22}$ or $\mathcal{R}_{23}$ and defect 1 in $\mathcal{R}_{33}$.

A special case to be noted is the situation in which $r = 0$, that is, r is identically zero. Recall that in this case we defined $\mu = -\infty$ and $\nu = 0$, so that r is said to have exact type $(-\infty, 0)$. The definition (24.1) remains in force in this case, so if $r = 0$, we say that r has defect $d = n$ in $\mathcal{R}_{mn}$, regardless of m.

The reason why defects matter has to do with the counting of zeros. Suppose $r = p/q \in \mathcal{R}_{mn}$ has exact type (μ, ν) and $\tilde{r} = \tilde{p}/\tilde{q}$ is another function in $\mathcal{R}_{mn}$. Then we have

$$r - \tilde{r} = \frac{p}{q} - \frac{\tilde{p}}{\tilde{q}} = \frac{p\tilde{q} - \tilde{p}q}{q\tilde{q}},$$

a rational function of type $(\max\{\mu + n, m + \nu\}, n + \nu)$. By (24.1), this implies that $r - \tilde{r}$ is of type $(m + n - d, 2n - d)$. Thus $r - \tilde{r}$ can have at most $m + n - d$ zeros, and this zero count is a key to equioscillation and uniqueness results.

Here is our main theorem. The study of rational best approximations goes back to Chebyshev [1859], including the idea of equioscillation, though without a precise statement of what form an alternant must take. Existence was first proved by de la Vallée Poussin [1911] and Walsh [1931], and equioscillation is due to Achieser [1930].

Theorem 24.1. Equioscillation characterization of best approximants.
A real function $f \in C([-1, 1])$ has a unique best approximation $r^* \in \mathcal{R}_{mn}^{\text{real}}$, and a function $r \in \mathcal{R}_{mn}^{\text{real}}$ is equal to r^* if and only if $f - r$ equioscillates between at least $m + n + 2 - d$ extreme points, where d is the defect of r in $\mathcal{R}_{mn}$.

"Equioscillation" here is defined just as in Chapter 10. For $f - r$ to equioscillate between k extreme points means that there exists a set of numbers $-1 \leq x_1 < \cdots < x_k \leq 1$ such that

$$f(x_j) - r(x_j) = (-1)^{j+i} \|f - r\|, \qquad 1 \leq j \leq k$$

with $i = 0$ or 1. Here and throughout this chapter, $\|\cdot\|$ is the supremum norm.

We now give some examples. To begin with, here is a function with a spike at $x = 0$:

```
x = chebfun('x'); f = exp(-100*x.^2);
```

Polynomial approximations of this function converge rather slowly. For example, it takes $n = 20$ to achieve one digit of accuracy:

```
[p,err] = remez(f,10);
subplot(1,2,1), hold off, plot(f-p), hold on
plot([-1 1],err*[1 1],'--k'), plot([-1 1],-err*[1 1],'--k')
FS = 'fontsize';
title('Error in type (10,0) approx',FS,9), ylim(.3*[-1 1])
[p,err] = remez(f,20);
subplot(1,2,2), hold off, plot(f-p), hold on
plot([-1 1],err*[1 1],'--k'), plot([-1 1],-err*[1 1],'--k')
title('Error in type (20,0) approx',FS,9), ylim(.3*[-1 1])
```
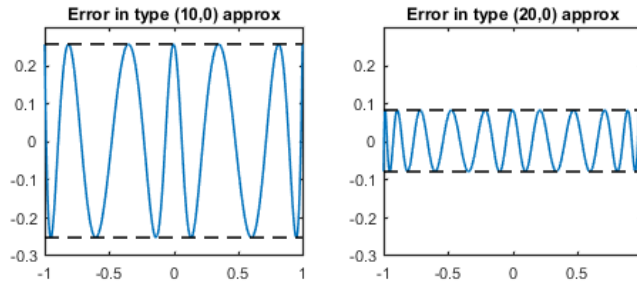
Notice that the extreme points of these error curves are distributed all across $[-1, 1]$, even though the challenging part of the function would appear to be in the middle. As discussed in Chapter 16, this is typical of polynomial best approximations.

If we switch to rational approximations, which can also be computed by Chebfun's `remez` command [Pachón & Trefethen 2009, Pachón 2010], the accuracy improves. Here we see error curves for approximations of types $(2, 2)$ and $(4, 4)$, with much smaller errors although the degrees are low. Note that most of the extreme points are now localized in the middle.

```
[p,q,rh,err] = remez(f,2,2);
subplot(1,2,1), hold off, plot(f-p./q), hold on
plot([-1 1],err*[1 1],'--k'), plot([-1 1],-err*[1 1],'--k')
title('Error in type (2,2) approx',FS,9), ylim(.1*[-1 1])
[p,q,rh,err] = remez(f,4,4);
subplot(1,2,2), hold off, plot(f-p./q), hold on
plot([-1 1],err*[1 1],'--k'), plot([-1 1],-err*[1 1],'--k')
title('Error in type (4,4) approx',FS,9), ylim(.1*[-1 1])
```
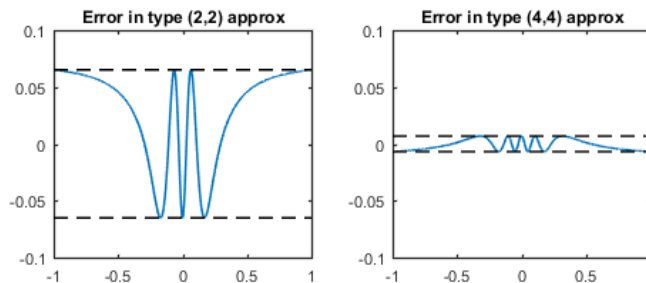

The error curves just plotted provide good examples of the role of the defect in the characterization of best approximants. The function f is even, and so are its best approximations (Exercise 24.1). Thus we expect that the type $(2, 2)$, $(3, 2)$, $(2, 3)$ and $(3, 3)$ best approximations will all be the same function, a rational function of exact type $(2, 2)$ whose error curve has 7 points of equioscillation. For $(m, n) = (2, 2)$, the defect is 0 and there is one more equioscillation point

than the minimum $m + n + 2 - d = 6$. For $(m, n) = (3, 2)$ or $(2, 3)$, the defect is 0 and the number of equiscillation points is exactly the minimum $m + n + 2 - d$. For $(m, n) = (3, 3)$, the defect is 1 and the number of equiscillation points is again exactly the minimum $m + n + 2 - d$.

Similarly, the error curve in the plot on the right, with 11 extrema, indicates that this rational function is a best approximation not only of type $(4, 4)$ but also of types $(5, 4)$, $(4, 5)$, and $(5, 5)$.

Here is another example, an odd function:

```
f = x.*exp(-5*abs(abs(x)-.3));
clf, plot(f), grid on, ylim(.4*[-1 1])
title('An odd function','fontsize',9)
```
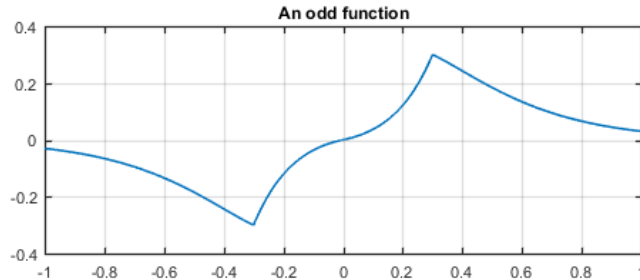

If we look for a best approximation of type $(4, 5)$, we find that the numerator has exact degree 3:

```
[p,q,rh,err] = remez(f,4,4); format short, chebpoly(p)
```

```
ans =
    0.0000    0.0093   -0.0000    0.0314   -0.0000
```
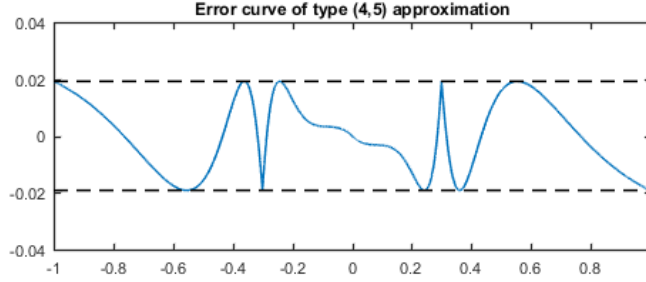
and the denominator has exact degree 4:

```
chebpoly(q)
```

```
ans =
    0.1174    0.0000    0.4065   -0.0000    0.3003
```

The defect is 1, so there must be at least $4 + 5 + 2 - 1 = 10$ extreme points in the error curve. In fact, there are exactly 10:

```
plot(f-p./q), hold on, ylim(.04*[-1 1])
plot([-1 1],err*[1 1],'--k'), plot([-1 1],-err*[1 1],'--k')
title('Error curve of type (4,5) approximation','fontsize',9)
```

Error curve of type (4,5) approximation

We conclude that r is the best approximation of types $(4,4), (4,5), (3,4)$ and $(3,5)$.

Let us now turn to the proof of Theorem 24.1. For polynomial approximations, our analogous theorem was Theorem 10.1, whose proof proceeded in four steps:

1. *Existence proof via compactness,*

2. *Equioscillation $\Rightarrow$ optimality,*

3. *Optimality $\Rightarrow$ equisoscillation,*

4. *Uniqueness proof via equioscillation.*

For rational functions, we shall follow the same sequence. The main novelty is in step 1, where compactness must be applied in a subtler way. We shall see an echo of this argument one more time in Chapter 27, in the proof of Theorem 27.1 for Padé approximants.

Part 1 of proof: Existence via compactness. For polynomial approximation, in Chapter 10, we noted that $\|f - p\|$ is a continuous function on $\mathcal{P}_n$, and since one candidate approximation was the zero polynomial, it was enough to look in the bounded subset $\{p \in \mathcal{P}_n : \|f - p\| \leq \|f\|\}$. Since this set was compact, the minimum was attained.

For rational functions, $\|f - r\|$ is again a continuous function on $\mathcal{R}_{mn}^{\mathrm{real}}$, and again it is enough to look in the bounded subset $\{r \in \mathcal{R}_{mn}^{\mathrm{real}} : \|f - r\| \leq \|f\|\}$, or more simply, the larger bounded set $\{r \in \mathcal{R}_{mn}^{\mathrm{real}} : \|r\| \leq 2\|f\|\}$. The difficulty is that bounded sets of rational functions are not in general compact. To illustrate this fact, consider the family of functions

$$r_\varepsilon(x) = \frac{x^3 + \varepsilon}{x^2 + \varepsilon}, \qquad (24.2)$$

where $\varepsilon > 0$ is a parameter. For each ε, $r_\varepsilon(x)$ is a continuous function on $[-1, 1]$ with $\|r_\varepsilon\| = 1$. As $\varepsilon \to 0$, however, r_ε behaves discontinuously:

$$\lim_{\varepsilon \to 0} r_\varepsilon(x) = \begin{cases} 1 & x = 0, \\ x & x \neq 0. \end{cases}$$

5

So we cannot find a limit function r_0 by taking a limit as $\varepsilon \to 0$. What saves us, however, is that the spaces of numerators and denominators are both compact, so we can argue that the numerators and denominators separately approach limits p_0 and q_0, which in this example would be x^3 and x^2. We then define a limiting rational function by $r_0 = p_0/q_0$ and argue by continuity that it has the necessary properties. This kind of reasoning is spelled out in greater generality in [Walsh 1931].

Suppose then that $\{r_k\}$ is a sequence of functions in $\mathcal{R}_{mn}^{\text{real}}$ with $\|r_k\| \leq 2\|f\|$ and
$$\lim_{k \to \infty} \|f - r_k\| = E = \inf_{r \in \mathcal{R}_{mn}^{\text{real}}} \|f - r\|.$$

Write each r_k in the form p_k/q_k with $p_k \in \mathcal{P}_m$, $q_k \in \mathcal{P}_n$, $q_k(x) \neq 0$ for all $x \in [-1, 1]$, and $\|q_k\| = 1$, hence $\|p_k\| \leq \|q_k\|\|r_k\| \leq 2\|f\|$. Since $\{p_k\}$ and $\{q_k\}$ lie in compact sets, we may assume by passing to a subsequence if necessary that $p_k \to p^*$ and $q_k \to q^*$ for some $p^* \in \mathcal{P}_m$ and $q^* \in \mathcal{P}_n$. Since $\|q_k\| = 1$ for each k, $\|q^*\| = 1$ too, and thus q^* is not identically zero but has at most a finite set of zeros on $[-1, 1]$. Now define $r^* = p^*/q^* \in \mathcal{R}_{mn}^{\text{real}}$. For all $x \in [-1, 1]$ except perhaps the zeros of q^*, $|f(x) - r^*(x)| = \lim_{k \to \infty} |f(x) - r_k(x)| \leq E$. By continuity, the same must hold for all $x \in [-1, 1]$, with p^* having zeros in $[-1, 1]$ wherever q^* does. Thus r^* is a best approximation to f. ∎

Part 2 of proof: Equioscillation $\Rightarrow$ optimality. Suppose $f - r$ takes equal extreme values with alternating signs at $m + n + 2 - d$ points $x_0 < x_1 < \cdots < x_{m+n+1-d}$, and suppose $\|f - \tilde{r}\| < \|f - r\|$ for some $\tilde{r} \in \mathcal{R}_{mn}^{\text{real}}$. Then $r - \tilde{r}$ must take nonzero values with alternating signs at the equioscillation points, implying that it must take the value zero in at least $m + n + 1 - d$ points in-between. However, as observed above, $r - \tilde{r}$ is of type $(m + n - d, 2n - d)$. Thus it cannot have $m + n + 1 - d$ zeros unless it is identically zero, a contradiction. ∎

Part 3 of proof: Optimality $\Rightarrow$ equioscillation. Suppose $f - r$ equioscillates at fewer than $m + n + 2 - d$ points, and set $E = \|f - r\|$. Without loss of generality suppose the leftmost extremum is one where $f - r$ takes the value $-E$. Then by a compactness argument, for all sufficiently small $\varepsilon > 0$, there are numbers $-1 < x_1 < \cdots < x_k < 1$ with $k \leq m + n - d$ such that $(f - r)(x) < E - \varepsilon$ for $x \in [-1, x_1 + \varepsilon] \cup [x_2 - \varepsilon, x_3 + \varepsilon] \cup [x_4 - \varepsilon, x_5 + \varepsilon] \cup \cdots$ and $(f - r)(x) > -E + \varepsilon$ for $x \in [x_1 - \varepsilon, x_2 + \varepsilon] \cup [x_3 - \varepsilon, x_4 + \varepsilon] \cup \cdots$. Let r be written in the form p/q, where p has degree $\mu \leq m - d$ and q has degree $\nu \leq n - d$, with p and q having no roots in common. The proof now consists of showing that r can be perturbed to a function $\tilde{r} = (p + \delta p)/(q + \delta q) \in \mathcal{R}_{mn}$ with the properties that $\|\tilde{r} - r\| < \varepsilon$ and $\tilde{r} - r$ is strictly negative for $x \in [-1, x_1 - \varepsilon] \cup [x_2 + \varepsilon, x_3 - \varepsilon] \cup [x_4 + \varepsilon, x_5 - \varepsilon] \cup \cdots$ and strictly positive for $x \in [x_1 + \varepsilon, x_2 - \varepsilon] \cup [x_3 + \varepsilon, x_4 - \varepsilon] \cup \cdots$. Such a function $\tilde{r}$ will have error less than E throughout the whole interval $[-1, 1]$. We calculate
$$\tilde{r} = \frac{p + \delta p}{q + \delta q} = \frac{(p + \delta p)(q - \delta q)}{q^2} + O(\|\delta q\|^2)$$

and therefore
$$\tilde{r} - r = \frac{q\delta p - p\delta q}{q^2} + O(\|\delta p\|\|\delta q\| + \|\delta q\|^2).$$

We are done if we can show that δp and δq can be chosen so that $q\delta p - p\delta q$ is a nonzero polynomial of degree exactly k with roots $x_1, \ldots, x_k$; by scaling this δp and δq sufficiently small, the quadratic terms above can be made arbitrarily small relative to the others, so that the required ε conditions are satisfied. This can be shown by the Fredholm alternative of linear algebra. The map from the $(m + n + 2)$-dimensional set of choices of δp and δq to the $(m + n + 1 - d)$-dimensional space of polynomials $q\delta p - p\delta q$ is linear. To show the map is surjective, it is enough to show that its kernel has dimension $d + 1$ but no more. Suppose then that $q\delta p - p\delta q$ is zero, that is, $q\delta p = p\delta q$. Then since p and q have no roots in common, all the roots of p must be roots of δp and all the roots of q must be roots of δq. In other words we must have $\delta p = gp$ and $\delta q = gq$ for some polynomial g. Since δp has degree no greater than m and δq has degree no greater than n, g can have degree no greater than d. The set of polynomials of degree d has dimension $d + 1$, so we are done. ∎

Part 4 of proof: Uniqueness via equioscillation. Finally, to prove uniqueness, suppose r is a best approximation whose error curve equioscillates between extreme points at $x_0 < x_1 < \cdots < x_{m+n+1-d}$, and suppose $\|f - \tilde{r}\| \leq \|f - r\|$ for some $\tilde{r} \in \mathcal{R}_{mn}^{\text{real}}$. Then (without loss of generality) $(r - \tilde{r})(x)$ must be ≤ 0 at $x_0, x_2, x_4, \ldots$ and ≥ 0 at $x_1, x_3, x_5, \ldots$. This implies that $r - \tilde{r}$ has roots in each of the $m + n + 1 - d$ closed intervals $[x_0, x_1], \ldots, [x_{m+n-d}, x_{m+n+1-d}]$, and since $r - \tilde{r}$ is a rational function of type $(m + n - d, 2n - d)$, the same must hold for its numerator polynomial. We wish to conclude that its numerator polynomial has at least $m + n + 1 - d$ roots in total, counted with multiplicity, implying that $r = \tilde{r}$. The argument for this is the same as given in the proof of Theorem 10.1. ∎

We have now finished the substantial mathematics. It is time to look at some of the consequences.

One of the recurring themes in the subject of rational approximation is the phenomenon of *square blocks in the Walsh table.* Suppose that a real function $f \in C([-1, 1])$ is given, and consider the set of all of its real rational best approximations of type (m, n) for various $m, n \geq 0$. We can imagine these laid out in an array, with m along the horizontal and n along the vertical. This array is called the *Walsh table* for f [Walsh 1934].

Generically, all the entries in the Walsh table for a given f will be distinct, and in this case we say that f is *normal.* Sometimes, however, certain entries in the table may be repeated, and in fact this is a frequent occurrence because it happens whenever f is even or odd. If f is even, then for any nonnegative integers j and k, all of its rational approximations of types $(2j, 2k)$, $(2j+1, 2k)$, $(2j, 2k + 1)$ and $(2j + 1, 2k + 1)$ must be the same. Similarly, if f is odd, then

all of its approximations of types $(2j + 1, 2k)$, $(2j + 2, 2k)$, $(2j + 1, 2k + 1)$ and $(2j + 2, 2k + 1)$ must be the same. We have already seen a number of examples.

More generally, repeated entries or "degeneracies" in the Walsh table may take complicated forms. Nevertheless the equioscillation condition imposes quite a bit of structure on the chaos. Degeneracies always appear precisely in a pattern of square blocks. The following statement of this result is taken from [Trefethen 1984], where a discussion of various aspects of this and related problems can be found. We shall return to the subject of square blocks in Chapter 27, on Padé approximation.

Theorem 24.2. Square blocks in the Walsh table. *The Walsh table of best real rational approximants to a real rational function $f \in C([-1, 1])$ breaks into precisely square blocks containing identical entries. (If f is rational, one of these will be infinite in extent.) The only exception is that if an entry $r = 0$ appears in the table, then it fills all of the columns to the left of some fixed index $m = m_0$.*

Proof. Given a nonrational function f, let $r \neq 0$ be a best approximation in $\mathcal{R}_{\mu\nu}^{\mathrm{real}}$ of exact type (μ, ν). (The cases of rational f or $r = 0$ can be handled separately.) By Theorem 24.1, the number of equioscillation points of $f - r$ is $\mu + \nu + 2 + k$ for some integer $k \geq 0$. We note that r is an approximation to f in $\mathcal{R}_{mn}^{\mathrm{real}}$ for any $m \geq \mu$ and $n \geq \nu$, and the defect is $\min\{m - \mu, n - \nu\}$. Thus by Theorem 24.1, r is the best approximation to f precisely for those values of (m, n) satisfying $m \geq \mu$, $n \geq \nu$, and $\mu + \nu + 2 + k \geq m + n + 2 - \min\{m - \mu, n - \nu\}$. The latter condition simplifies to $n \leq \nu + k$ and $m \leq \mu + k$, showing that r is the best approximation to f precisely in the square block $\mu \leq m \leq \mu + k$, $\nu \leq n \leq \nu + k$. ∎

Within a square block in the Walsh table, the defect d is equal to zero precisely in the first column and the first row. An approximation with $d = 0$ is sometimes said to be *nondegenerate*. It can have more points of equioscillation than the generic number $m + n + 2$, but never fewer.

As mentioned above, the theory of equioscillation and degeneracies is very appealing mathematically. As an example we note a result due to Werner [1964], in completion of earlier work of Maehly and Witzgall [1960]: the type (m, n) *best approximation operator*, which maps functions f to their best approximations r_{mn}^*, is continuous at f with respect to the supremum norm if and only if $f \in \mathcal{R}_{mn}$ or the corresponding function r_{mn}^* is nondegenerate. The essential reason for this effect is that if a function r^* is the best approximation to f in a nontrivial square block, then a small perturbation $f \to \tilde{f}$ might fracture that block into pieces of size 1×1 [Trefethen 1984]. If (m, n) corresponds to a degenerate position in the block, with $d > 0$, then the best approximation $\tilde{r}^*$ for such an $\tilde{f}$ would need to have a higher equioscillation number than that of r^* for f, requiring $\tilde{r}^*$ to be far from r^* if $\|f - r^*\|$ is positive.

These complications hint at some of the practical difficulties of rational approximation. For example, the Remez algorithm is based on explicit manipulation of alternant sets. If the number of extremal points is not known a priori, it is plausible that one may expect numerical difficulties in certain circumstances. Indeed, this is the case, and so far as I am aware, no implementation of the Remez algorithm for rational approximation, including Chebfun's, can be called fully robust. Other kinds of algorithms may have better prospects.

We finish by returning to the matter of best *complex* approximations to real functions. Nonuniqueness of certain complex rational approximations was pointed out by Walsh in the 1930s. Later Lungu [1971] noticed, following a suggestion of Gonchar, that the nonuniqueness arises even for approximation of a real function f on $[-1, 1]$, with examples as simple as type $(1, 1)$ approximation of $|x|$. (Exercise 24.3 gives another proof that there must exist such examples.) These observations were rediscovered independently by Saff and Varga [1978a]. Ruttan [1981] showed that complex best approximations are always better than real ones in the strict lower-right triangle of a square block, that is, when a type (m, n) best approximation equioscillates in no more than $m + n + 1$ points. Trefethen and Gutknecht [1983a] showed that for every (m, n) with $n \geq m + 3$, examples exist where the ratio of the optimal complex and real errors is arbitrarily small. Levin, Ruttan and Varga showed that the minimal ratio is exactly $1/3$ for $n = m + 2$ and exactly $1/2$ for $1 \leq n \leq m + 1$ [Ruttan & Varga 1989]. None of this has much to do with practical approximation, but it is fascinating.

SUMMARY OF CHAPTER 24. *Any real function $f \in C([-1, 1])$ has a unique best approximation $r^* \in \mathcal{R}_{mn}^{\text{real}}$ with respect to the ∞-norm, and r^* is characterized by having an error curve that equioscillates between at least $m + n + 2 - d$ extreme points, where d is the defect of r in $\mathcal{R}_{mn}$. In the Walsh table of all best approximations to f indexed by m and n, repeated entries, if any, lie in exactly square blocks.*

Exercise 24.1. Approximating even functions. Prove that if a real function $f \in C([-1, 1])$ is even, then its real best approximations of all types (m, n) are even.

Exercise 24.2. Approximating the Gaussian. The first figures of this chapter considered lower degree polynomial and rational approximations of $\exp(-100x^2)$ on $[-1, 1]$. Make a plot of the errors in approximations of types $(n, 0)$ and (n, n), now taking n as high as you can. (You may find that the `cf` command takes you farther than `remez`.) How do the polynomial and rational approximations compare?

Exercise 24.3. Complex approximations and nonuniqueness. (a) Suppose a real function $f \in C([-1, 1])$ takes both the values 1 and -1. Prove that no real rational function $r \in \mathcal{R}_{0n}^{\text{real}}$, for any n, can have $\|f - r\| < 1$. (b) On the other hand, show that for any $\varepsilon > 0$, there is a complex rational function $r \in \mathcal{R}_{0n}$ for some n with $\|f - r\| < \varepsilon$. (Hint: perturb f by an imaginary constant and consider its reciprocal.) (c) Conclude that type $(0, n)$ complex rational best approximations in $C([-1, 1])$ are nonunique in general for large enough n.

Exercise 24.4. A function with a spike. Plot chebfuns of the function (24.2) for $\varepsilon = 1, 0.1, \ldots, 10^{-6}$ and determine the polynomial degree $n(\varepsilon)$ of the chebfun in each case. What is the observed asymptotic behavior of $n(\varepsilon)$ as $\varepsilon \to 0$? How accurately can you explain this observation based on the theory of Chapter 8?

Exercise 24.5. de la Vallée Poussin lower bound. Suppose an approximation $r \in \mathcal{R}_{mn}^{\text{real}}$ to $f \in C([-1, 1])$ approximately equioscillates in the sense that there are points $-1 \le s_0 < s_1 < \cdots < s_{m+n+1-d} \le 1$ at which $f - r$ alternates in sign with $|f(s_j) - r(s_j)| \ge \varepsilon$ for some $\varepsilon > 0$, where d is the defect of r in $\mathcal{R}_{mn}$. Show that the best approximation $r^* \in \mathcal{R}_{mn}^{\text{real}}$ satisfies $\|f - r^*\| \ge \varepsilon$. (Compare Exercise 10.3.)

Exercise 24.6. A rational lethargy theorem. Let $\{\varepsilon_n\}$ be a sequence decreasing monotonically to 0. Adapt the proof of Exercise 10.7 to show that that there is a function $f \in C([-1, 1])$ such that $\|f - r_{nn}^*\| \ge \varepsilon_n$ for all n.