SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Talet e och dess historia

av

Aleyna Bulduk

2025 - No L5

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

Talet e och dess historia

Aleyna Bulduk

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Håkan Granath

2025

Sammanfattning

Talet e, eller det så kallade Eulers tal, är en av de mest kända talen i matematiken. Elever möter detta tal oftast i anknytning till exponentialfunktioner och logaritmer och får lära sig att talet är basen till den naturliga logaritmen ln. Utöver det ges ofta inte särskilt mycket fakta om talets uppkomst och egenskaper, vilket i regel skapar kunskapsluckor bland studenter. Mot den bakgrunden är syftet med arbetet att bidra till en fördjupad förståelse av talet e. I detta arbete ska vi därmed utforska talets historia, bevisa att det är ett irrationellt och till och med ett transcendent tal, härleda dess kedjebråk, samt diskutera och jämföra några av talets approximationsmetoder, däribland gränsvärde, serie och kedjebråk.

Abstract

The number e, otherwise known as Euler's number, is one of the most famous numbers in mathematics. Students encounter this number usually in the context of exponential functions and logarithms, in which they acquire that the number is the base of the natural logarithm ln. Besides that, rarely any facts regarding its origin and properties are given, which, in general, creates gaps in knowledge among students. Against this background, this paper aims to contribute to a deeper understanding of the number e. In this study, we will thus explore the history of the number, prove that it is an irrational and even a transcendental number, derive its continued fraction expansion, as well as discuss and compare some of the number's approximation methods, including limit, series, and continued fraction.

Acknowledgements

I would like to express my sincere gratitude and appreciation to my supervisor Håkan Granath for his continuous support and guidance during the writing of this paper. Also, I want to thank my family for always being by my side and encouraging me to become the best version of myself.

Innehåll

1 Inledning

1.1 Bakgrund

Talet e, eller det så kallade Eulers tal, är många bekant med från matematikkurserna på gymnasiet. I dessa kontexter presenteras talet vanligtvis i samband med exponentialfunktioner och logaritmer där det brukar benämnas som basen för den naturliga logaritmen ln och uppfyller följande samband

$$\ln y = x \Leftrightarrow y = e^x.$$

Med andra ord innebär detta att den naturliga logaritmen av ett tal är exponenten för vilket e behöver upphöjas till för att vara lika med talet [1]. Dessa beräkningar görs numera enkelt med en räknare där man kan knappa in e utan att ha minsta lilla tanke på vad det faktiskt står för. Samtidigt ges lite utrymme för fakta om talet i undervisningen, vilket medför att det saknas omfattande insikter om talet bland studenter. Alltså finns det ofta en kunskapslucka bland elever och studenter vad gäller talets uppkomst och egenskaper.

1.2 Syfte och disposition

Syftet med detta arbete är därmed att fördjupa förståelsen av talet e genom att beskriva dess historia, bevisa att det är ett irrationellt och transcendent tal, härleda dess kedjebråk, samt redogöra för och jämföra några av de metoder som finns för att approximera värdet av talet, däribland serie, gränsvärde, och kedjebråk. Arbetet kommer följa motsvarande upplägg, först ska talets historia presenteras, därefter bevis om talets irrationalitet respektive transcendens, härledning av talets kedjebråk, och slutligen approximationsmetoder.

2 Historien om talet e

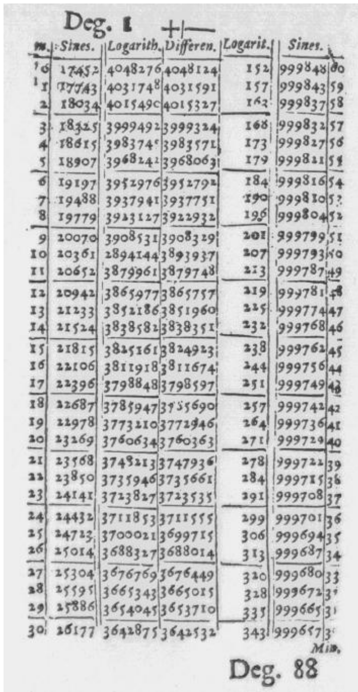
I detta avsnitt ska vi undersöka historien om talet e genom att använda källorna [3] och [2].

Överraskande nog upptäcktes talet e inte i samband med logaritmer även om studierna som bedrevs kring logaritmer var ganska nära till denna upptäckt. År 1614 publicerade den skotske matematikern John Napier en tabell på logaritmer av diverse tal i sitt verk *Mirifici logarithmorum canonis descriptio*. Det är oklart hur Napier exakt kom på dessa logaritmer men han var känd för att vara kunnig inom geometri och aritmetik. Innan Napiers tid studerade man geometriska talföljder och fann ett samband mellan termerna i en geometrisk talföljd och de motsvarande exponenterna av kvoten. En geometrisk talföljd är en följd av tal där två på varandra följande tal har en konstant kvot/proportion. Exempelvis är $1, q, q^2, q^3, ...$ en geometrisk följd vars konstanta kvot är q. Den tyske matematikern Micheal Stifel redogörde för sambandet i sin bok *Arithmetica integra* (1544) och formulerade potenslagarna som används än idag. Vi har $q^m \cdot q^n = q^{m+n}$ där multiplikation av två termer i en geometrisk talföljd är ekvivalent med addition av deras exponenter. Därtill har vi även $q^m/q^n = q^{m-n}$ där division av två termer i en geometrisk talföljd är ekvivalent med subtraktion av deras exponenter. Vidare har vi att $q^{-n} = 1/q^n$ och $q^0 = 1$. Dessa samband var grundtanken bakom logaritmerna. Ifall vilket positivt tal som helst skrivs som en potens av ett givet tal (som senare kallas för en bas), så är multiplikation och division av olika tal ekvivalent med att addera respektive subtrahera deras exponenter. Vidare om talet multipliceras med sig själv n gånger innebär detta att dess exponent multipliceras med n, och för att finna den n:e roten av ett tal delas dess exponent med n.

Napier använde sig av dessa förhållanden för att skapa logaritmer och förenkla numeriska beräkningar. Napier ville med sina logaritmer handskas med både heltal och reella tal. För att logaritmerna skulle fylla de flesta luckorna på tallinjen kämpade han med att finna en bas som är så nära 1 som möjligt. Efter flertal år kom han slutligen fram till basen $1 - 10^{-7}$. Därefter spenderade han 20 år för att finna termerna som skulle förekomma i hans logaritmtabell. Notera att Napier åstadkom detta med endast penna och papper! Dessa tabeller fick namnet logaritm eftersom Napier kallade exponenten av varje potens för ordet logaritm. I sin första logaritmtabell använde han formeln $N = 10^7(1 - 10^{-7})^L$, där exponenten L är Napiers logaritm av N, för att beräkna varje logaritm. Eftersom basen $1 - 10^{-7}$ är mindre än 1 innebär detta att Napiers logaritmer minskar i takt med större tal, jämfört med dagens logaritmer till basen 10 som ökar.

Napier var egentligen ganska nära att upptäcka talet $1/e$ och i sin tur e med sina logaritmer. Ifall vi dividerar variablerna N och L i hans definition ovan med 10^7

får vi ekvationen $M = ((1-10^{-7})^{10^7})^K$ där $M = N/10^7$ och $K = L/10^7$. Talet $(1-10^{-7})^{10^7}$ är mycket nära talet $1/e$, alltså är Napiers logaritmer i princip logaritmer med basen $1/e$. Dessvärre uppmärksammades inte denna bas eftersom på den tiden uppfattades inte logaritmer på samma sätt som det görs idag. Nuförtiden tänker vi på logaritmer som inversa funktioner till exponentialfunktioner. Förr i tiden användes logaritmer däremot endast som ett verktyg för att förenkla numeriska beräkningar. Därav fanns det förmodligen inte ett stort behov av att ta reda på en sådan bas då. Apropå detta utvecklades konceptet kring logaritmbaser huvudsakligen när logaritmbasen 10 introducerades.



Deg. 1 +|—

m.	Sines.	Logarith.	Differen.	Logarit.	Sines.	
0	17452	4048276	4048124	152	999848	60
1	17743	4031748	4031591	157	999843	59
2	18034	4015490	4015327	163	999837	58
3	18325	3999492	3999324	168	999832	57
4	18615	3983745	3983571	173	999827	56
5	18907	3968242	3968063	179	999821	55
6	19197	3952976	3952792	184	999816	54
7	19488	3937941	3937751	190	999810	53
8	19779	3923127	3922932	196	999804	52
9	20070	3908531	3908329	201	999799	51
10	20361	2894144	3893937	207	999793	50
11	20652	3879961	3879748	213	999787	49
12	20942	3865977	3865757	219	999781	48
13	21233	3852186	3851960	225	999774	47
14	21524	3838582	3838351	232	999768	46
15	21815	3825161	3824923	238	999762	45
16	22106	3811918	3811674	244	999756	44
17	22396	3798848	3798597	251	999749	43
18	22687	3785947	3755690	257	999742	42
19	22978	3773210	3772946	264	999736	41
20	23269	3760634	3760363	271	999729	40
21	23568	3748213	3747936	278	999722	39
22	23850	3735946	3735661	284	999715	38
23	24141	3723827	3723535	291	999708	37
24	24432	3711853	3711555	299	999701	36
25	24723	3700021	3699715	306	999694	35
26	25014	3688327	3688014	313	999687	34
27	25304	3676769	3676449	320	999680	33
28	25595	3665343	3665015	328	999672	3[illegible]
29	25886	3654045	3653710	335	999665	3[illegible]
30	26177	3642875	3642532	343	999657	3[illegible]

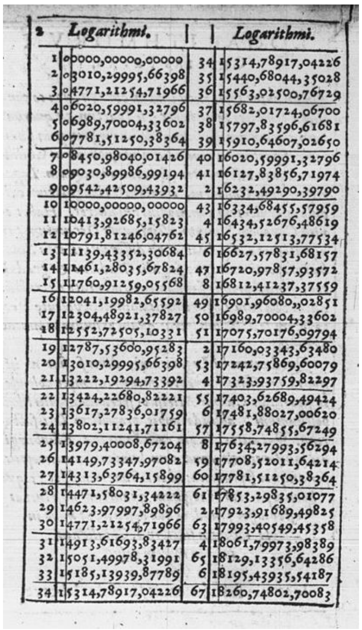
Min.

Deg. 88

Figur 1: En sida av Napiers logaritmtabeller.

Napiers logaritmtabeller fick ett stort genomslag inom naturvetenskapen. Forskare runtom i Europa och även Kina tog till sig denna upptäckt mycket snabbt och använde det i sina beräkningar. En av de första som använde sig av logaritmer var den tyske astronomen Johannes Kepler som tillämpade logaritmer i sina beräkningar av planetarnas banor. Därtill blev den engelske matematikern Henry Briggs så pass fascinerad av upptäckten att han åkte till Skottland för att träffa Napier. Tillsammans med Napier utvecklade Briggs dagens logaritmer med basen 10. De enades om att $\log 10 = 1$. Alltså kunde numera ett positivt tal N skrivas med basen 10 enligt formeln $N = 10^L$ där L kallas för Briggs logaritm av N, eller med andra ord $L = \log N$.

Vartefter publicerade Briggs tabeller av logaritmer med basen 10 i sitt verk *Arithmetica logarithmica* (1624). Briggs logaritmtabeller innehöll tiologaritmer av heltalen 1–20000 och 90000–100000. Tiologaritmerna av heltalen 20000–90000 beräknades därefter av den nederländske matematikern Adriaan Vlacq och resultatet återfinns i den andra upplagan av *Arithmetica logarithmica* (1628).



2	Logarithmi.		Logarithmi.
1	00000,00000,00000	34	15314,78917,04226
2	03010,29995,66398	35	15440,68044,35028
3	04771,21254,71966	36	15563,02500,76729
4	06020,59991,32796	37	15682,01724,06700
5	06989,70004,33602	38	15797,83596,61681
6	07781,51250,38364	39	15910,64607,02650
7	08450,98040,01426	40	16020,59991,32796
8	09030,89986,99194	41	16127,83856,71974
9	09542,42509,43932	2	16232,49290,39790
10	10000,00000,00000	43	16334,68455,57959
11	10413,92685,15823	4	16434,52676,48619
12	10791,81246,04762	45	16532,12513,77534
13	11139,43352,30684	6	16627,57831,68157
14	11461,28035,67824	47	16720,97857,93572
15	11760,91259,05568	8	16812,41237,37559
16	12041,19982,65592	49	16901,96080,,02851
17	12304,48921,37827	50	16989,70004,33602
18	12552,72505,10331	51	17075,70176,09794
19	12787,53600,95283	2	17160,03343,63480
20	13010,29995,66398	53	17242,75869,60079
21	13222,19294,73392	4	17323,93759,82297
22	13424,22680,82221	55	17403,62689,49424
23	13617,27836,01759	6	17481,88027,00620
24	13802,11241,71161	57	17558,74855,67249
25	13979,40008,67204	8	17634,27993,56294
26	14149,73347,97082	59	17708,52011,64214
27	14313,63764,15899	60	17781,51250,38364
28	14471,58031,34222	61	17853,29835,01077
29	14623,97997,89896	2	17923,91689,49825
30	14771,21254,71966	63	17993,40549,45358
31	14913,61693,83427	4	18061,79973,98389
32	15051,49978,31991	65	18129,13356,64286
33	15185,13939,87789	6	18195,43935,54187
34	15314,78917,04226	67	18260,74802,70083

Figur 2: En sida av Briggs första tabell över tiologaritmer, publicerad i *Logarithmorum Chilias Prima* (1617). Sidan visar tiologaritmerna av heltalen 0-67.

Vartefter bedrevs ett flertal studier i detta fält, vilket intresserade kan läsa om i [3] och [2], men en annan som var nära nog att upptäcka e genom logaritmer var den nederländske matematikern Christiaan Huygens. År 1661 undersökte han den rektangulära hyperbeln $y = \frac{1}{x}$ och fann ett samband mellan arean under grafen och logaritmen. Arean av en sådan graf är integralen av funktionen $\frac{1}{x}$, alltså logaritmen $\ln x$. I detta fall är arean mellan 1 och talet e lika med 1, det vill säga $\ln e = 1$, men dessvärre identifierades talet e inte heller i detta sammanhang.

Faktum är att talet e upptäcktes i samband med studier kring sammansatt ränta, så kallad ränta på ränta. Som namnet tyder handlar sammansatt ränta om att tjäna ränta på både det investerade kapitalbeloppet och tidigare ränta. Exempelvis en investering på 500 kronor med en årsränta på 2 procent kommer återinvesteras och efter 3 år växa till $500 \cdot (1,02)^3 \approx 531$ kronor. Formeln som används är alltså

$$S = P(1 + r)^t, \tag{2.1}$$

där P står för det investerade kapitalbeloppet från början, r för årsräntesatsen uttryckt i decimalform, t för tid i år och S för summan. Däremot är det mycket vanligt att det sammansatta räntan beräknas per halvår, kvartal, månad, vecka, eller dag i finansiella institutioner. Ifall den beräknas n gånger per år kommer räntan r delas upp i n delar för att beräkna räntan för varje period. Då förvandlas ekvationen (2.1) till

$$S = P\left(1 + \frac{r}{n}\right)^n, \tag{2.2}$$

där $\frac{r}{n}$ är räntan per period och n är antal perioder. Antag nu att $P = 1$ och $r = 1$, vilket innebär att 1 krona investeras med en generös räntesats på 100 procent, därmed blir formeln

$$S = \left(1 + \frac{1}{n}\right)^n. \tag{2.3}$$

Vad kommer nu hända med formeln (2.3) om n stiger och går mot oändligheten? Vi har alltså gränsvärdet

$$\lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n. \tag{2.4}$$

Detta gränsvärde är faktiskt viktigt eftersom det är definitionen för talet e som upptäcktes av den schweiziske matematikern Jacob Bernoulli år 1683. Detta innebär att e är faktiskt det första talet som definierades av ett gränsvärde! Man kan undra hur det kommer sig att e upptäcktes inom ramen för ekonomi, särskilt sammansatt ränta. Dock bör vi ha i åtanke att 1600-talet var en period där internationell handel florerade och sammansatt ränta var av stor vikt, därför är det inte så förvånande att talet e identifierades under denna tidsperiod.

Även om det inte var den schweiziske matematikern Leonhard Euler som upptäckte talet e är notationen e, som används än idag, etablerad av honom och användes för första gången år 1731 i hans brev till den preussiska matematikern Christian Goldbach. Varefter har han gjort ett flertal studier och upptäckter kring talet. I sin bok *Introductio in Analysin infinitorum* (1748) bevisar han att talet e är gränsvärdet (2.4) och även att det är summan av den oändliga serien

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots \tag{2.5}$$

Dessutom härleder han kedjebråk för talet e och finner ett mönster, vilket kommer presenteras mer ingående i följande avsnitt. Utöver detta bevisade Euler att talet e är irrationellt, varefter den franske matematikern Charles Hermite bevisade år 1873 att det även är transcendent.

3 Irrationalitet

I detta avsnitt ska vi gå igenom begreppet irrationalitet och upptäckten av de första irrationella talen. Avsnittet återger fakta från boken [3].

Rationella tal, alltså tal som kan skrivas som ett bråk av två heltal, har varit känd för mänskligheten i flera tusentals år och dess historia sträcker sig ända till de forntida civilisationerna. De forntida egyptierna och babylonierna uppfann många olika sätt att approximera rationella tal och räkna med dem. Däremot var det pythagoréerna, alltså antikens greker som hanförde sig till matematikern Pythagoras och hans lära, som högaktade rationella tal och utsåg det till kärnan i deras matematik och filosofi. Efter att ha experimenterat med musik och upptäckt skala och intervaller som kan definieras med rationella tal ansåg de att detsamma måste gälla för allting på jorden. Alltså var de övertygade om att allt kunde beskrivas med rationella tal och därmed dominerade detta deras världsbild i hög grad. Men dessvärre varade inte denna åskådning särskilt länge. Detta eftersom de stötte på ett problem när de skulle beräkna längden av diagonalen på en enhetskvadrat, alltså en kvadrat med sidor som är 1 längdenhet. I detta fall ger Pythagoras sats att $x^2 = 1^2 + 1^2$, där x är längden av diagonalen. Lite förenkling ger att $x = \sqrt{2}$ och detta tal försökte de skriva i bråkform. Dock kom de fram till slutsatsen att $\sqrt{2}$ inte kan skrivas som ett bråk av två heltal och därmed upptäcktes det första irrationella talet. Denna upptäckt var däremot mycket förbluffande för pythagoréerna, då de var inställda på att alla tal är rationella. De kunde approximera talet $\sqrt{2}$ men ändå inte lyckas skriva det som ett rationellt tal. Därmed valde de att förneka talet och hemlighålla denna storslagna upptäckt. Enligt en sägen bestämde sig dock en av pythagoréerna vid namn Hippasus att ta bladet från munnen och offentliggöra det, men detta slutade med att han kastades i havet av sina följeslagare. Så småningom spreds kunskapen om irrationella tal vidare och vartefter hittades fler irrationella tal, såsom kvadratroten av varenda primtal.

De rationella talen tillsammans med de irrationella talen utgör de reella talen. Ett reellt tal är ett tal som kan skrivas i decimalform. Till skillnad från de rationella talen som har antingen en ändlig, eller oändlig och repeterande (periodisk) decimalutveckling har de irrationella talen en oändlig och icke-repeterande (icke-periodisk) decimalutveckling. Med det menas att decimalerna fortsätter i all oändlighet och inte upprepas i samma mönster. Exempelvis är $\frac{1}{5} = 0.2$ respektive $\frac{1}{3} = 0.3333..$ exempel på två rationella tal medan 2.7182... är exempel på ett irrationellt tal. Den sistnämnda är decimalutvecklingen av talet e. Trots att irrationella tal inte kan skrivas som ett bråk av två heltal kan man approximera värdet av dessa tal med valfritt antal decimalers nogrannhet med hjälp av rationella tal. Exempelvis är $\frac{27}{10} = 2.7$, $\frac{1359}{500} = 2.718$, och $\frac{271828}{100000} = 2.71828$ alla approximationer av talet e med stigande noggrannhet.

4 Kedjebråk

I detta avsnitt ska vi gå igenom grunderna i kedjebråk som ges i källorna [7], [8] och [9].

Kedjebråk används bland annat för att approximera reella tal med rationella tal. Det är en viss form att skriva ett tal och i synnerhet gestaltas det som

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{a_3 + \cfrac{b_4}{a_4 + ... + \cfrac{b_{n-1}}{a_{n-1} + \cfrac{b_n}{a_n}}}}}} \tag{4.1}$$

Bokstäverna b_n där $n \geq 1$ betecknar täljarna, medan a_n för alla $n \geq 1$ representerar nämnarna i kedjebråksutvecklingen. Samtliga tal är heltal och kan antingen vara negativa eller positiva beroende på talets värde. Ifall $b_n = 1$ för alla $n \geq 1$ benämns kedjebråket som regelbundet. Insättning av $b_n = 1$ i (4.1) ger därmed ett regelbundet kedjebråk

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + ... + \cfrac{1}{a_{n-1} + \cfrac{1}{a_n}}}}}} \tag{4.2}$$

I detta arbete kommer vi begränsa oss till regelbundna kedjebråk där samtliga a_n för $n \geq 0$ är positiva heltal. För fler typer av kedjebråk hänvisas läsaren till [7] som är en engelsk översättning av Eulers *De Fractionibus Continuis Dissertatio*. En kedjebråksutveckling av ett tal a kan även kortfattat skrivas som $[a_0; \ldots, a_n]$ där talen a_k för $k \in [0, n]$ kallas för partiella kvoter av kedjebråket.

4.1 Kedjebråk av rationella tal

Varenda rationellt tal har ett ändligt kedjebråk enligt följande sats

Sats 4.3. *[9, Sats 1.1] Varje ändligt kedjebråk representerar ett rationellt tal. Omvänt, varje rationellt tal $\frac{a}{b}$ kan representeras som ett ändligt kedjebråk.*

Bevis. Det första påståendet är uppenbart, då vi kan alltid räkna baklänges och omvandla en ändlig kedjebråksutveckling till ett rationellt tal. Låt oss bevisa att varje rationellt tal har ett ändligt kedjebråk. Antag att $\frac{a}{b}$ för $b \geq 1$ är ett rationellt tal. Successiv division ger

$$\begin{aligned}
\frac{a}{b} &= a_0 + \frac{r_0}{b} \quad \text{där } 0 < r_0 < b, \\
\frac{b}{r_0} &= a_1 + \frac{r_1}{r_0} \quad \text{där } 0 < r_1 < r_0, \\
\frac{r_0}{r_1} &= a_2 + \frac{r_2}{r_1} \quad \text{där } 0 < r_2 < r_1, \\
&\vdots \\
\frac{r_{n-3}}{r_{n-2}} &= a_{n-1} + \frac{r_{n-1}}{r_{n-2}} \quad \text{där } 0 < r_{n-1} < r_{n-2}, \\
\frac{r_{n-2}}{r_{n-1}} &= a_n + \frac{r_n}{r_{n-1}} = a_n \quad \text{där } r_n = 0.
\end{aligned} \tag{4.4}$$

Efter ett ändligt antal steg kommer vi erhålla resten $r_n = 0$ och divisionen kommer således att upphöra. Ifall vi sätter ihop dessa ekvationer kommer vi få

$$\begin{aligned}
\frac{a}{b} &= a_0 + \frac{r_0}{b} = a_0 + \cfrac{1}{\cfrac{b}{r_0}} = a_0 + \cfrac{1}{a_1 + \cfrac{r_1}{r_0}} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cfrac{r_0}{r_1}}} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{r_2}{r_1}}} = \dots \\
&= a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + ... + \cfrac{1}{a_{n-1} + \cfrac{r_{n-1}}{r_{n-2}}}}} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + ... + \cfrac{1}{a_{n-1} + \cfrac{1}{\cfrac{r_{n-2}}{r_{n-1}}}}}} \\
&= a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + ... + \cfrac{1}{a_{n-1} + \cfrac{1}{a_n}}}}
\end{aligned}$$

Alltså kommer vi få den ändliga och regelbundna kedjebråksutvecklingen av $\frac{a}{b}$. $\square$

Låt oss ta ett exempel. Succesiv division av 49 med 29 ger

$$\begin{aligned}
\frac{49}{29} &= 1 + \frac{20}{29}, \\
\frac{29}{20} &= 1 + \frac{9}{20}, \\
\frac{20}{9} &= 2 + \frac{2}{9}, \\
\frac{9}{2} &= 4 + \frac{1}{2}, \\
\frac{2}{1} &= 2.
\end{aligned}$$

Vi får alltså de partiella kvoterna $a_0 = 1, a_1 = 1, a_2 = 2, a_3 = 4, a_4 = 2$ och kan därmed skriva kedjebråksutveckingen

$$\frac{49}{29} = 1 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{4 + \cfrac{1}{2}}}} = [1; 1, 2, 4, 2]. \tag{4.5}$$

Kedjebråk grundar sig på Euklides algoritm. Euklides algoritm används för att hitta den största gemensamma delaren, SGD, av två heltal. Låt oss syngliggöra sambandet mellan kedjebråk och Euklides algoritm som i [9]. Ekvationerna i (4.4) är faktiskt samma ekvationer som fås genom att tillämpa Euklides algoritm på a och b

$$\begin{aligned}
a &= a_0 b + r_0 \quad \text{där } 0 < r_0 < b, \\
b &= a_1 r_0 + r_1 \quad \text{där } 0 < r_1 < r_0, \\
r_0 &= a_2 r_1 + r_2 \quad \text{där } 0 < r_2 < r_1, \\
&\vdots \\
r_{n-3} &= a_{n-1} r_{n-2} + r_{n-1} \quad \text{där } 0 < r_{n-1} < r_{n-2}, \\
r_{n-2} &= a_n r_{n-1} + r_n = a_n r_{n-1} \quad \text{där } r_n = 0.
\end{aligned} \tag{4.6}$$

Notera att ekvationerna i (4.6) är en omskrivning av ekvationerna i (4.4). Exempelvis fås den första ekvationen i (4.6) genom att multiplicera den första ekvationen i (4.4) med nämnaren b, och så vidare.

Följande är några exempel på kedjebråksutvecklingar av olika rationella tal.

$$\frac{115}{82} = 1 + \frac{33}{82} = 1 + \frac{1}{\frac{82}{33}} = 1 + \frac{1}{2 + \frac{16}{33}} = 1 + \frac{1}{2 + \frac{1}{\frac{33}{16}}} = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{16}}} = [1; 2, 2, 16]. \tag{4.7}$$

$$\frac{63}{17} = \cdots = 3 + \frac{1}{1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2}}}} = [3; 1, 2, 2, 2]. \tag{4.8}$$

4.2 Kedjebråk av reella tal

Vi går vidare med kedjebråksalgoritmen för reella tal som finns i [8]. Denna kedjebråksalgoritm är en mer generell version av kedjebråksalgoritmen (4.4) eftersom den gäller för både rationella och irrationella tal. Antag att $a \in \mathbb{R}^+$ och definiera två talföljder $a_0, a_1, \ldots,$ och $\xi_1, \xi_2, \ldots$ genom

$$\begin{aligned} &a_0 = [a] \quad a = a_0 + \frac{1}{\xi_1} \\ &\text{och rekursivt som} \\ &a_i = [\xi_i] \quad \xi_i = a_i + \frac{1}{\xi_{i+1}} \quad \text{för } i \geq 1. \end{aligned} \tag{4.9}$$

I detta sammanhang betecknar $[x]$ det största heltalet som är mindre än eller lika med $x \in \mathbb{R}^+$. Genom att tillämpa formlerna ovan får vi kedjebråksutvecklingen av talet a, alltså

$$a = a_0 + \frac{1}{\xi_1} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\xi_2}} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\xi_3}}} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\xi_4}}}} = \dots$$

$$= a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + ... + \cfrac{1}{a_{n-1} + \cfrac{1}{\xi_n}}}}}}$$

Ifall a är ett rationellt tal kommer kedjebråksalgoritmen för rationella tal (4.4) och kedjebråksalgoritmen för reella tal (4.9) vara likadana. Med det sagt kommer ξ_i så småningom att bli ett heltal som i kedjebråksalgoritmen (4.4) och därmed kommer kedjebråksutvecklingen att upphöra. Däremot har irrationella tal oändliga kedjebråksutvecklingar. Med andra ord kan vi skriva kedjebråksutvecklingen som $[a_0; a_1, a_2, \dots, a_n, \dots]$ där punkterna betyder att processen fortskrider.

Ett exempel på ett irrationellt tal är gyllene snittet $\gamma = \frac{1+\sqrt{5}}{2}$, som uppfyller ekvationen $\gamma^2 = \gamma + 1$. Denna ekvation kan i sin tur skrivas som $\gamma = 1 + \frac{1}{\gamma}$ där $1 = [\gamma]$. Kedjebråksalgoritmen (4.9) ger att

$$\gamma = 1 + \frac{1}{\gamma} = 1 + \cfrac{1}{1 + \cfrac{1}{\gamma}} = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{\gamma}}} = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \dots}}}} = [1; 1, 1, 1, 1, \dots].$$

4.3 Konvergenter

Som tidigare nämnt är kedjebråksutvecklingar av rationella tal ändliga och upphör efter ett antal steg, till skillnad från kedjebråksutvecklingar av irrationella tal som är oändliga. Däremot kan man approximera värdet av ett irrationellt tal med rationella tal så nära som möjligt. Faktum är att om man stegvis trunkerar kedjebråksutvecklingen av ett irrationellt tal kommer man erhålla en följd av rationella tal som konvergerar mot det önskade irrationella talet. Dessa rationella tal kallas för konvergenter och är de optimala approximationerna av ett givet tal, vilket vi kommer återkomma till och förtydliga ytterligare inom ramen för irrationella tal nedan. Apropå detta kan man få en bättre approximation ju längre kedjebråksutvecklingen blir innan det trunkeras.

För att tydliggöra vad konvergenter är kan vi ta ett rationellt tal som exempel. Ifall kedjebråksutvecklingen av talet $\frac{115}{82}$ från (4.7) stegvis trunkeras kommer vi få konvergenterna

$$1, \quad 1+\frac{1}{2}=\frac{3}{2}, \quad 1+\cfrac{1}{2+\cfrac{1}{2}}=\frac{7}{5}, \quad 1+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{16}}}=\frac{115}{82}.$$

Formellt kan konvergenter skrivas som

$$\begin{aligned}
c_0 &= \frac{a_0}{1} = \frac{p_0}{q_0}, \\
c_1 &= a_0 + \frac{1}{a_1} = \frac{a_0 a_1 + 1}{a_1} = \frac{p_1}{q_1}, \\
&\vdots \\
c_k &= a_0 + \cfrac{1}{a_1 + \ldots + \cfrac{1}{a_{k-1} + \cfrac{1}{a_k}}} = \frac{p_k}{q_k},
\end{aligned}$$

där $c_k = [a_0; a_1, \ldots, a_k] = \frac{p_k}{q_k}$, för $k = 0, \ldots, n$ kallas för den k:e konvergenten av kedjebråket $[a_0; a_1, \ldots, a_n]$ för ett tal a [9]. Vi kan finna den k:e konvergenten till en given kedjebråksutveckling genom följande sats.

Sats 4.10. *[9, Sats 1.3] Täljaren p_k och nämnaren q_k till den k:e konvergenten c_k av kedjebråket $[a_0, a_1, \ldots, a_n]$ ges av de rekursiva talföljderna*

$$\begin{aligned}
p_k &= a_k p_{k-1} + p_{k-2} \text{ för } k \geq 2 \\
q_k &= a_k q_{k-1} + q_{k-2} \text{ för } k \geq 2
\end{aligned}$$

med startvärdena

$$\begin{aligned}
p_0 &= a_0 \quad p_1 = a_1 a_0 + 1 \\
q_0 &= 1 \quad q_1 = a_1.
\end{aligned}$$

Vi tar kedjebråket av $\frac{63}{17} = [3; 1, 2, 2, 2]$ från (4.8) som exempel och finner dess konvergenter genom att tillämpa satsen ovan.

$$
\begin{aligned}
c_0 &= \frac{p_0}{q_0} = \frac{a_0}{1} = a_0 = 3,\\
c_1 &= \frac{p_1}{q_1} = \frac{a_0 a_1 + 1}{a_1} = \frac{3 \cdot 1 + 1}{1} = 4,\\
c_2 &= \frac{p_2}{q_2} = \frac{a_2 p_1 + p_0}{a_2 q_1 + q_0} = \frac{2 \cdot 4 + 3}{2 \cdot 1 + 1} = \frac{11}{3},\\
c_3 &= \frac{p_3}{q_3} = \frac{a_3 p_2 + p_1}{a_3 q_2 + q_1} = \frac{2 \cdot 11 + 4}{2 \cdot 3 + 1} = \frac{26}{7},\\
c_4 &= \frac{p_4}{q_4} = \frac{a_4 p_3 + p_2}{a_4 q_3 + q_2} = \frac{2 \cdot 26 + 11}{2 \cdot 7 + 3} = \frac{63}{17}.
\end{aligned}
$$

I båda exemplen med $\frac{115}{82}$ och $\frac{63}{17}$ har konvergenterna alternerande större och mindre värden än talets egentliga värde. Konvergenter med udda index har ett värde som är större än talet själv, medan de med jämn index har ett mindre värde. Exempelvis är 3 och $\frac{11}{3}$ mindre än $\frac{63}{17}$, medan 4 och $\frac{26}{7}$ är större. Däremot kommer samtliga konvergenter att så småningom konvergera mot det egentliga värdet, vilket kommer beskrivas mer ingående snart.

Konvergenter $c_n = \frac{p_n}{q_n}$ till oändliga kedjebråk definieras på liknande sätt som tidigare 4.10 med samma rekursiva formler och startvärden för p_n och q_n. Täljaren p_n och nämnaren q_n uppfyller följande samband som finns i [8]

$$p_{n+1} q_n - p_n q_{n+1} = (-1)^n \text{ för } n \geq 0. \tag{4.11}$$

Detta samband gäller både för ändliga och oändliga kedjebråk. Division med $q_n q_{n+1}$ på båda sidor av (4.11) ger

$$\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n} = \frac{(-1)^n}{q_n q_{n+1}}.$$

Vi vet att $c_n = \frac{p_n}{c_n}$. Substitution ger därmed

$$c_{n+1} - c_n = \frac{(-1)^n}{q_n q_{n+1}}. \tag{4.12}$$

Dessutom har vi i [8] att

$$c_{n+2} - c_n = \frac{(-1)^n (q_{n+2} - q_n)}{q_n q_{n+1} q_{n+2}}. \tag{4.13}$$

Låt oss bevisa detta samband. Vi får

$$\begin{aligned}
c_{n+2} - c_n &= c_{n+2} - c_{n+1} + c_{n+1} - c_n \\
&= \frac{p_{n+2}}{q_{n+2}} - \frac{p_{n+1}}{q_{n+1}} + \frac{(-1)^n}{q_n q_{n+1}} \\
&= \frac{p_{n+2}q_{n+1} - p_{n+1}q_{n+2}}{q_{n+1}q_{n+2}} + \frac{(-1)^n(q_{n+2})}{q_n q_{n+1} q_{n+2}}.
\end{aligned}$$

Notera att uttrycket $p_{n+2}q_{n+1} - p_{n+1}q_{n+2} = (-1)^{n+1}$ då n byts ut mot $n+1$ i sambandet (4.11). Vi får alltså

$$\begin{aligned}
c_{n+2} - c_n &= \frac{(-1)^{n+1}}{q_{n+1}q_{n+2}} + \frac{(-1)^n(q_{n+2})}{q_{n+1}q_{n+2}} \\
&= \frac{q_n(-1)^{n+1}}{q_n q_{n+1} q_{n+2}} + \frac{(-1)^n(q_{n+2})}{q_n q_{n+1} q_{n+2}} \\
&= \frac{(-1)^n(-q_n) + (-1)^n(q_{n+2})}{q_n q_{n+1} q_{n+2}} \\
&= \frac{(-1)^n(q_{n+2} - q_n)}{q_n q_{n+1} q_{n+2}}.
\end{aligned}$$

Därmed har vi bevisat sambandet (4.13).

Dessa två samband synliggör hur konvergenterna c_n av ett kedjebråk förändras när n ökar och förhållandet mellan konvergenter med udda respektive jämn index. Sambanden (4.12) och (4.13) ger följande som framgår i [9]

$$c_0 < c_2 < c_4 < \cdots < c_{2n} < \cdots < c_{2n+1} < \cdots < c_5 < c_3 < c_1 \tag{4.14}$$

Detta kan sammanfattas med följande sats

Sats 4.15. *[9, Sats 3.3] Konvergenterna med udda index c_{2n+1} av ett oändligt och regelbundet kedjebråk skapar en avtagande följd, medan konvergenterna med jämn index c_{2n} skapar en växande följd. Dessutom är varje konvergent med jämn index mindre än varenda konvergent med udda index.*

Vidare har konvergenter ett gränsvärde. Detta eftersom varje växande och uppåt begränsad, och tvärtom, varje avtagande och nedåt begränsad talföljd är konvergent [10]. Konvergenter med udda index är avtagande och nedåt begränsad då de konvergerar mot ett värde som är större än samtliga konvergenter med jämn index. Å andra sidan är konvergenter med jämn index växande och uppåt begränsad för att de konvergerar mot ett värde som är mindre än alla konvergenter med udda index. Med andra ord existerar gränsvärdena $\lim_{n\to\infty} c_{2n+1}$ och $\lim_{n\to\infty} c_{2n}$. Låt oss återge beviset i [9, s.67] om att dessa gränsvärden är lika, alltså $\lim_{n\to\infty} c_{2n+1} = \lim_{n\to\infty} c_{2n}$.

Ersätt n med $2t$ och $n+1$ med $2t+1$ i (4.12). Vi får

$$c_{2t+1} - c_{2t} = \frac{(-1)^{2t}}{q_{2t}q_{2t+1}}.$$

Eftersom $(-1)^{2t} = 1$ för samtliga värden på t ger detta

$$c_{2t+1} - c_{2t} = \frac{1}{q_{2t}q_{2t+1}}.$$

När t går mot oändligheten kommer nämnaren $q_{2t}q_{2t+1}$ i högerledet att växa, eftersom termerna i de rekursiva talföljderna i 4.10 är positiva heltal och kommer att stiga. Detta medför att högerledet $\frac{1}{q_{2t}q_{2t+1}}$ närmar sig 0. Alltså kommer skillnaden $c_{2t+1} - c_{2t}$ att närma sig 0 när t går mot ändligheten. Med andra ord är $\lim_{t\to\infty} c_{2t+1} = \lim_{t\to\infty} c_{2t}$. Detta gränsvärde är i självfallet det irrationella talet, vilket beskrivs i följande sats

Sats 4.16. *[9, Sats 3.6] Om ett irrationellt tal x uttrycks som ett oändligt och regelbundet kedjebråk $[a_0; a_1, \ldots, a_n, \ldots]$ med hjälp av kedjebråksalgoritmen* (4.9) *så kommer konvergenterna $c_0, c_1, \ldots, c_n, \ldots$ att konvergera mot talet x.*

Konvergenter i en kedjebråksutveckling av ett irrationellt tal är faktiskt de bästa rationella approximationerna av talet eftersom inget annat rationellt tal med mindre nämnare kan approximera värdet bättre, enligt satsen nedan

Sats 4.17. *[8, Sats 2] Låt $\alpha > 0$ vara ett irrationellt tal och antag att $\frac{p}{q}$ är ett rationellt tal sådant att $|\alpha - \frac{p}{q}| < \frac{1}{2q^2}$. Då är $\frac{p}{q}$ en konvergent i kedjebråksutvecklingen av α.*

5 Talet e är irrationellt

5.1 Fouriers bevis

Den franske matematikern Joseph Fourier bevisade att e är ett irrationellt tal genom att använda serien av e (2.5) som Euler hade bevisat, alltså

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots$$

Vi kan konstatera att denna serie är positiv eftersom alla dess termer är strikt större än 0.

Det framgår i [5] att beviset inte publicerades av Fourier själv, utan av Janot de Stainville i boken *Mélanges d'analyse algébrique et de géométrie* (1815). De Stainville menade att han i sin tur har erfarit beviset genom en annan fransk matematiker vid namn Louis Poinsot som har personligen fått beviset av Fourier. Beviset är ett motsägelsebevis som innebär att man först gör ett antagande att e är ett rationellt tal men slutligen finner en motsägelse. Jämfört med Eulers och Hermites bevis som tillämpar kedjebråk är detta bevis relativt lättförståeligt och effektivt att skriva. Innan vi går vidare till att bevisa att e är ett irrationellt tal är det angeläget att redovisa satsen kring jämförelsekriterium för konvergens hos positiva serier eftersom denna sats utnyttjas i beviset.

Sats 5.1. *[4, Sats 2] Jämförelsekriteriet: Antag att* $0 \leq a_k \leq b_k$ *för alla* $k =$ *0,1,2,..., då gäller att* $\sum_{k=0}^{\infty} a_k$ *konvergerar om* $\sum_{k=0}^{\infty} b_k$ *konvergerar.*

Vi börjar med att fastställa att e är ett tal mellan 2 och 3.

Sats 5.2. *Talet e är större än 2 men mindre än 3.*

Bevis. [5, s.9] Eftersom summan av de två första termerna i serien (2.5) är lika med 2 och de resterande termerna är positiva kan vi fastställa att e har ett värde som är större än 2. Vidare är summan av de resterande termerna mindre än 1, vilket kan bevisas genom jämförelsekriteriet. Eftersom $n! \geq 2^{n-1}$ för $n \geq 2$ så gäller det att

$$\frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots < \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots$$

Termerna i den högra sidan kommer från den kända serien $\sum_{n=1}^{\infty} \frac{1}{2^n}$ som konvergerar mot talet 1. Därmed är summan av de resterande termerna $\frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots < 1$, vilket ger att $e < 3$. Alltså har vi visat att $2 < e < 3$. $\square$

Låt oss nu gå vidare till Fouriers bevis om att e är ett irrationellt tal. Följande bevis är en kombination av bevisen i [5] och [6].

Sats 5.3. *Talet e är irrationellt.*

Bevis. Som tidigare nämnt är beviset ett motsägelsebevis som innebär att vi först antar att e är ett rationellt tal, alltså att e kan skrivas som ett bråk av två heltal. Antag nu att $a, b \in \mathbb{Z}^+$ och $b > 1$ så att

$$e = \frac{a}{b}.$$

Vidare har vi e som serien (2.5), alltså kan vi skriva e som

$$\frac{a}{b} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots$$

Multiplikation med heltalet $b!$ ger att

$$(b-1)!a = b!\left(1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{b!} + \frac{1}{(b+1)!} + \frac{1}{(b+2)!} + \dots\right).$$

Vänsterledet $(b-1)!a$ är ett heltal då produkten av två heltal är ett heltal. Därmed ska vi kontrollera om högerledet också är ett heltal. Högerledet kan delas upp så att

$$(b-1)!a = b!\left(1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{b!}\right) + b!\left(\frac{1}{(b+1)!} + \frac{1}{(b+2)!} + \dots\right).$$

Vi kan observera att produkten av $b!$ och delsumman $S_b = \sum_{n=0}^{b} \frac{1}{n!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots + \frac{1}{b!}$ blir ett heltal eftersom $b!$ är delbar med samtliga nämnare. Eftersom vänsterledet och produkten $b!S_b$ är båda heltal måste även resten, alltså produkten av $b!$ och delsumman $R_b = \sum_{k=1}^{\infty} \frac{1}{(b+n)!} = \frac{1}{(b+1)!} + \frac{1}{(b+2)!} + \frac{1}{(b+3)!} + \dots$ vara ett heltal, men vi kommer snart finna en motsägelse. Notera att $b!R_b$ är större än 0 eftersom samtliga termer är positiva. En omskrivning av ekvationen ger

$$0 < (b-1)!a - b! \cdot S_b = b! \cdot R_b.$$

Förenkla högerledet $b! \cdot R_b$ så vi får

$$b! \cdot R_b = \frac{1}{b+1} + \frac{1}{(b+1)(b+2)} + \frac{1}{(b+1)(b+2)(b+3)} + \dots$$

Uttrycket ovan har ett lägre värde än serien $\sum_{k=1}^{\infty} \frac{1}{(b+1)^k} = \frac{1}{b+1} + \frac{1}{(b+1)^2} + \frac{1}{(b+1)^3} + \ldots$, alltså kan vi använda jämförelsekriteriet 5.1 även här, vilket ger

$$\begin{aligned}
0 &< \frac{1}{b+1} + \frac{1}{(b+1)(b+2)} + \frac{1}{(b+1)(b+2)(b+3)} + \ldots \\
&< \frac{1}{b+1} + \frac{1}{(b+1)^2} + \frac{1}{(b+1)^3} + \ldots \\
&= \frac{1}{b+1}\left(1 + \frac{1}{b+1} + \frac{1}{(b+1)^2} + \ldots\right) \\
&= \frac{1}{b+1}\sum_{k=0}^{\infty}\left(\frac{1}{b+1}\right)^k \\
&= \frac{1}{b+1}\left(\frac{1}{1-\frac{1}{b+1}}\right) \\
&= \frac{1}{b}.
\end{aligned} \tag{5.4}$$

I vilket serien $\sum_{k=0}^{\infty}\left(\frac{1}{b+1}\right)^k$ är geometrisk och kan därmed skrivas som uttrycket $\frac{1}{1-\frac{1}{b+1}}$ [10, Sats 9, s.178].

Eftersom b är ett positivt heltal som är större än 1 innebär detta att $1/b < 1$, därmed

$$0 < \frac{1}{b+1} + \frac{1}{(b+1)(b+2)} + \frac{1}{(b+1)(b+2)(b+3)} + \cdots < 1.$$

Däremot existerar inget heltal mellan 0 och 1, alltså har vi faktiskt kommit fram till en motsägelse. Med det sagt är vårt ursprungliga antagande om att e är ett rationellt tal falskt. Talet e är därav ett irrationellt tal. $\square$

5.2 Eulers bevis och upptäckt av kedjebråket för *e*

I detta avsnitt ska vi kort redogöra för Eulers bevis av irrationalitet av talet e. Detta avsnitt grundar sig på artikel [11].

Euler bevisade år 1748 att e är irrationellt i sin bok *Introductio in Analysin infinitorum*. Beviset är icke-sammanhängande eftersom Euler hoppar över flera steg och hävdar istället att sina slutsatser gäller. Med anledning av detta ska vi endast få en liten inblick på beviset, men intresserade hänvisas till [11]. Euler bevisade att e är ett irrationellt tal genom att bevisa att dess kedjebråksutveckling är oändlig, alltså på formen $[a_0; a_1, \ldots, a_n, \ldots]$. Euler använder talet $\frac{e-1}{2} \approx 0,8591409142295 = \frac{8591409142295}{10000000000000}$ som exempel [11]. Låt oss istället använda talet i decimalform och tillämpa kedjebråksalgoritmen (4.9) för att konstruera kedjebråksutvecklingen av talet

$$
\begin{aligned}
\frac{e-1}{2} &\approx 0,8591409142295 = 0 + \cfrac{1}{1 + 0,16395341374} \\
&= 0 + \cfrac{1}{1 + \cfrac{1}{6 + 0,09929355656}} = 0 + \cfrac{1}{1 + \cfrac{1}{6 + \cfrac{1}{10 + 0,071146957}}} \\
&= 0 + \cfrac{1}{1 + \cfrac{1}{6 + \cfrac{1}{10 + \cfrac{1}{14 + 0,0554149069}}}} \\
&= 0 + \cfrac{1}{1 + \cfrac{1}{6 + \cfrac{1}{10 + \cfrac{1}{14 + \cfrac{1}{18 + 0,0456858261}}}}} = \cdots = [0; 1, 6, 10, 14, 18, \ldots].
\end{aligned}
$$

Ifall värdet av e avrundades mer precist i början menar Euler att följden av de partiella kvoterna följer mönstret 1,6,10,14,18,22,26,34,. . . , där de partiella kvoterna a_n för $n \geq 2$ skapar en aritmetisk talföljd i vilket varje nästkommande term är summan av den föregående termen och 4. Alltså ökar varje nämnare med 4 efter nämnaren 6. Eftersom nämnarna ökar och inte slutar är kedjebråksutvecklingen oändlig. Med andra ord kan talet $\dfrac{e-1}{2}$ inte vara rationell. Då talet är irrationell kan e inte vara rationell heller.

På motsvarande sätt härledde Euler den oändliga kedjebråksutvecklingen av e

$$e = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{4 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{6 + \dots}}}}}}}} = [2; 1, 2, 1, 1, 4, 1, 1, 6, \dots]. \quad (5.5)$$

Denna kedjebråksutveckling har han härlett från en differentialekvation så kallad Ricattis ekvation. Ekvationen ser ut som följande

$$ady + y^2dx = x^{\frac{-4n}{2n+1}}.$$

Vi kommer inte att gå igenom hur Euler har härlett denna differentialekvation och hur den leder till kedjebråksutvecklingen (5.5), men intresserade läsare hänvisas till [11].

6 Hermites härledning av kedjebråket för e

I detta avsnitt ska vi gå igenom Hermites härledning av kedjebråksutvecklingen av e. Avsnittet grundar sig på artikeln [12] som presenterar en kort och direkt version av Hermites bevis av kedjebråksutvecklingen (5.5). Detta bevis uppkom som ett resultat av Hermites bevis av transcendens av talet e år 1873. Vi kommer bidra med ytterligare förklaringar och beräkningar till beviset i artikeln [12].

Låt oss omformulera kedjebråksutvecklingen (5.5) till $[1; 0, 1, 1, 2, 1, 1, 4, 1, 1, 6, \dots]$ genom att ersätta den första partiella kvoten 2 med $[1; 0, 1]$. Dessa två kedjebråksutvecklingar är ekvivalenta eftersom

$$e = 1 + \cfrac{1}{0 + \cfrac{1}{1 + \cfrac{1}{z}}} = 2 + \frac{1}{z}$$

där $z = [1; 2, 1, 1, 4, 1, 1, 6, \dots]$.

Antag nu att $[a; a_1, a_2, \dots]$ är kedjebråksutvecklingen $[1; 0, 1, 1, 2, 1, 1, 4, 1, 1, 6, \dots]$. Vi ska bevisa att den k:e konvergenten $c_k = p_k/q_k$ av e konvergerar mot e, alltså

$$\lim_{k \to \infty} \frac{p_k}{q_k} = e,$$

där p_k och q_k definieras på motsvarande sätt som i 4.10.

De partiella kvoterna följer mönstret

$$\begin{aligned} a_{3k+1} &= 2k \text{ för } k \geq 0, \\ a_{3k} = a_{3k+2} &= 1 \text{ för } k \geq 0. \end{aligned}$$

Enligt sats 4.10 definieras p_k och q_k med de rekursiva formlerna

$$\begin{array}{r|ll} p_{3n} = p_{3n-1} + p_{3n-2} & q_{3n} = q_{3n-1} + q_{3n-2} & \text{för } n \geq 1, \\ p_{3n+1} = 2np_{3n} + p_{3n-1} & q_{3n+1} = 2nq_{3n} + q_{3n-1} & \text{för } n \geq 1, \\ p_{3n+2} = p_{3n+1} + p_{3n} & q_{3n+2} = q_{3n+1} + q_{3n} & \text{för } n \geq 0, \end{array} \tag{6.1}$$

samt med startvärdena $p_0 = 1$, $q_0 = 1$, $p_1 = 1$ och $q_1 = 0$. Notera att konvergenten $c_1 = p_1/q_1$ är odefinierad eftersom $q_1 = 0$, men detta är inget bekymmer i härledningen av kedjebråksutvecklingen. Låt oss definiera integralerna

$$\begin{aligned} A_n &= \int_0^1 \frac{x^n(x-1)^n}{n!} e^x dx \\ B_n &= \int_0^1 \frac{x^{n+1}(x-1)^n}{n!} e^x dx \\ C_n &= \int_0^1 \frac{x^n(x-1)^{n+1}}{n!} e^x dx \end{aligned} \tag{6.2}$$

Följande sats visar sambandet mellan integralerna och konvergenternas täljare p_k och nämnare q_k

Sats 6.3. *[12, Proposition 1] För $n \geq 0$ gäller det att $A_n = q_{3n}e - p_{3n}, B_n = p_{3n+1} - q_{3n+1}e$ och $C_n = p_{3n+2} - q_{3n+2}e$.*

Bevis. Vi har att

$$\begin{aligned}
A_0 &= \int_0^1 e^x dx = e - 1 = q_0 e - p_0, \\
B_0 &= \int_0^1 xe^x dx = 1 = p_1 - q_1, \\
C_0 &= \int_0^1 (x-1)e^x dx = 2 - e = a_2 p_1 + p_0 - (a_2 q_1 + q_0)e = p_2 - q_2 e.
\end{aligned}$$

Vidare ska vi bevisa de rekursiva formlerna

$$A_n = -B_{n-1} - C_{n-1}, \tag{6.4}$$

$$B_n = -2nA_n + C_{n-1}, \tag{6.5}$$

$$C_n = B_n - A_n. \tag{6.6}$$

Låt oss börja med att bevisa den tredje rekursiva formeln (6.6).

$$B_n - A_n = \int_0^1 \frac{x^n(x-1)^n}{n!} e^x (x-1) dx = \int_0^1 \frac{x^n(x-1)^{n+1}}{n!} e^x dx = C_n.$$

För att bevisa (6.4) alltså $A_n + B_{n-1} + C_{n-1} = 0$ använder vi produktregeln, vilket ger oss

$$\frac{d}{dx}\left(\frac{x^n(x-1)^n}{n!}e^x\right) = \frac{x^n(x-1)^n}{n!}e^x + \frac{x^n(x-1)^{n-1}}{(n-1)!}e^x + \frac{x^{n-1}(x-1)^n}{(n-1)!}e^x.$$

Integrera båda sidor av ekvationen ovan, därmed

$$\left[\frac{x^n(x-1)^n}{n!}e^x\right]_0^1 = \int_0^1 \frac{x^n(x-1)^n}{n!}e^x dx + \int_0^1 \frac{x^n(x-1)^{n-1}}{(n-1)!}e^x dx + \int_0^1 \frac{x^{n-1}(x-1)^n}{(n-1)!}e^x dx$$

som i sin tur kan skrivas som $0 = A_n + B_{n-1} + C_{n-1}$. Alltså har vi fått den rekursiva formeln (6.4).

Vi går vidare till att bevisa (6.5) alltså $B_n + 2nA_n - C_{n-1} = 0$ genom att använda produktregeln på motsvarande sätt som tidigare

$$\begin{aligned}
\frac{d}{dx}\left(\frac{x^n(x-1)^{n+1}}{n!}e^x\right) &= \frac{x^n(x-1)^{n+1}}{n!}e^x + (n+1)\frac{x^n(x-1)^n}{n!}e^x + \frac{x^{n-1}(x-1)^{n+1}}{(n-1)!}e^x \\
&= (x-1)\frac{x^n(x-1)^n}{n!}e^x + (n+1)\frac{x^n(x-1)^n}{n!}e^x + (x-1)\frac{x^{n-1}(x-1)^n}{(n-1)!}e^x \\
&= \frac{x^{n+1}(x-1)^n}{n!}e^x + 2n\frac{x^n(x-1)^n}{n!}e^x - \frac{x^{n-1}(x-1)^n}{(n-1)!}e^x.
\end{aligned}$$

Integrera vänster och högerledet av ekvationen, vilket ger

$$\left[\frac{x^n(x-1)^{n+1}}{n!}e^x\right]_0^1 = \int_0^1 \frac{x^{n+1}(x-1)^n}{n!}e^x dx + 2n\int_0^1 \frac{x^n(x-1)^n}{n!}e^x dx - \int_0^1 \frac{x^{n-1}(x-1)^n}{(n-1)!}e^x dx.$$

Detta kan därmed skrivas som $0 = B_n + 2nA_n - C_{n-1}$.

Alltså har vi bevisat de rekursiva formlerna för A_n, B_n och C_n. Det följer från de rekursiva formlerna (6.1) att motsvarande formler uppfylls av $a_n = q_{3n}e - p_{3n}$, $b_n = p_{3n+1} - q_{3n+1}e$ och $c_n = p_{3n+2} - q_{3n+2}e$ från sats 6.3. Alltså har vi två uppsättningar av talföljder som har samma rekursion och samma startvärde, vilket innebär att de är lika. Därmed $a_n = A_n$, $b_n = B_n$, och $c_n = C_n$, som gäller för $n \geq 0$. □

Låt oss nu bevisa Eulers kedjebråksutveckling (5.5) av e.

Sats 6.7. *[12, Sats 1]* $e = [2; 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10, 1, \ldots]$.

Bevis. Variabeln x i integralerna A_n, B_n och C_n från (6.2) ligger mellan 0 och 1, och $n \geq 0$. Vidare är integranderna i dessa integraler mindre än eller lika med $e/n!$ för $x \in [0, 1]$. När n går mot oändligheten kommer dessa integraler att därmed gå mot 0 på grund av nämnaren $n!$. Vi har att $q_k \geq 1$ för $k \geq 2$ och av sats 6.3 följer det alltså att

$$\lim_{n\to\infty} \frac{p_{3n}}{q_{3n}} = \lim_{n\to\infty} e - \frac{A_n}{q_{3n}} = e,$$
$$\lim_{n\to\infty} \frac{p_{3n+1}}{q_{3n+1}} = \lim_{n\to\infty} e - \frac{B_n}{q_{3n+1}} = e,$$
$$\lim_{n\to\infty} \frac{p_{3n+2}}{q_{3n+2}} = \lim_{n\to\infty} e - \frac{C_n}{q_{3n+2}} = e.$$

Med andra ord konvergerar konvergenterna mot e

$$\lim_{k\to\infty} \frac{p_k}{q_k} = e = [2; 1, 2, 1, 1, 4, \ldots].$$

Därmed har vi bevisat kedjebråksutvecklingen (5.5). □

7 Transcendens

I detta avsnitt ska vi gå igenom begreppet transcendens och upptäckten av de första transcendenta talen. Avsnittet återger fakta från boken [3].

Drygt 2500 år efter pythagoréernas upptäckt av det första irrationella talet $\sqrt{2}$ upptäcktes en ny typ av tal, de så kallade transcendenta talen. I själva verket är de flesta talen vi möter inom matematiken lösningar till ekvationer, eller rättare sagt polynomekvationer med heltalskoefficienter. Vi har exempelvis talen $5, -\frac{7}{2}$ och $\sqrt{2}$ som utgör lösningar till polynomekvationerna $x-5=0$, $7x+2=0$, respektive $x^2-2=0$. Även om vi i detta avsnitt ska behandla de reella talen kan vi kort nämna att även det imaginära talet $i=\sqrt{-1}$ är en lösning till polynomekvationen $3x^2+3=0$. De reella tal som uppfyller polynomekvationer med heltalskoefficienter kallas för algebraiska tal. I den bemärkelsen är varje rationellt tal $\frac{a}{b}$ algebraisk eftersom den uppfyller ekvationen $bx-a=0$. Ifall ett tal inte är algebraisk så måste den vara irrationell. Däremot gäller inte det omvända, alltså kan ett irrationellt tal vara algebraisk liksom talet $\sqrt{2}$ som är både irrationell och algebraisk.

I början av 1800-talet grubblade matematiker över existensen av reella tal som är icke-algebraiska, eller så kallade transcendenta tal, men sådana tal verkade faktiskt vara svåra att hitta. Dock dröjde det inte länge innan det första transcendenta talet upptäcktes. År 1844 bevisade den franske matematikern Joseph Liouville förekomsten av icke-algebraiska tal. Med hjälp av sitt bevis kunde han konstruera flera transcendenta tal. Ett av de talen i fråga är $\sum_1^\infty \frac{1}{10^{n!}}$. Detta tal fick namnet Liouvilles nummer, där värdet av varje term avtar kraftigt på grund av exponenten $n!$ av nämnaren 10. Ett annat icke-algebraisk tal är 0,123456789101112... vars decimaler är följden av de naturliga talen.

Som nämnt tidigare i avsnittet om irrationalitet upptäcktes irrationella tal i samband med ett geometrisk problem. Transcendenta tal, å andra sidan, snarare konstruerades med syftet att bevisa deras existens. Alltså var de konstgjorda tal. Så småningom riktades uppmärksamheten på e som var känd att vara irrationell sedan drygt ett sekel bakåt. Man började undra om talet även är transcendent. Liouville bevisade att e inte uppfyller andragradsekvationer med heltalskoefficienter, men beviset är inte tillräckligt för att visa att talet är transcendent. Detta eftersom det krävs att e inte uppfyller någon polynomekvation alls. År 1873 bevisade den franske matematikern Charles Hermite att e är ett transcendent tal.

Vi avslutar detta avsnitt med en intressant upptäckt. År 1874 upptäckte den tyske matematikern Georg Cantor att det finns fler irrationella tal än rationella, men även att det faktiskt finns fler transcendenta tal än algebraiska. Med detta menas att de flesta reella tal är irrationella, i vilket de flesta irrationella tal är transcendentala.

8 Bevis av att talet e är transcendent

I detta avsnitt ska vi titta på beviset av att e är ett transcendent tal. Följande bevis är kombination av bevisen i [13], [14] och [15]. Beviset i fråga är ett motsägelsebevis, då vi först antar att e är transcendent, alltså att talet uppfyller polynomekvationer med heltalskoefficienter, men därefter når en motsägelse.

I beviset ska vi tillämpa gammafunktionen [16], vilket är en matematisk funktion som generaliserar fakulteten $n!$ för $n \in \mathbb{N}$. Den definieras som

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt,$$

där $x > 0$. För $x = n + 1$ får vi att

$$\Gamma(n+1) = \int_0^\infty t^n e^{-t} dx = n! \tag{8.1}$$

Sats 8.2. *e är ett transcendent tal.*

Bevis. Anta nu att e är ett algebraiskt tal, alltså rot till en grad n polynomekvation $p(x)$ med icke-noll heltalskoefficienter. Vi har därmed

$$p(e) = a_n e^n + \cdots + a_1 e + a_0 = 0, \tag{8.3}$$

där $a_0, a_1, \ldots, a_n \in \mathbb{Z}$ och $a_0 \neq 0$, eftersom ifall $a_0 = 0$ så kan vi alltid erhålla en ny polynomekvation $p(x) = 0$ sådan att $p(e) = 0$. För att bevisa att (8.3) inte gäller ska vi visa att potenser av e kan skrivas som

$$e^k = \frac{N_k + \delta_k}{N}, \quad k = 1, 2, \ldots, n, \tag{8.4}$$

där termerna N och N_k är heltal medan δ_k är mycket små tal. De definieras som

$$N = \frac{1}{(p-1)!} \int_0^\infty e^{-x} f(x) dx, \tag{8.5}$$

$$N_k = \frac{1}{(p-1)!} \int_k^\infty e^{k-x} f(x) dx, \tag{8.6}$$

$$\delta_k = \frac{1}{(p-1)!} \int_0^k e^{k-x} f(x) dx, \tag{8.7}$$

i vilket $f(x) = x^{p-1}((x-1)(x-2)\ldots(x-n))^p$ och p är ett godtyckligt primtal som uppfyller $p > n$ och $p > |a_0|$. Detta medför att $p \nmid n$ och $p \nmid a_0$. I slutet av beviset kommer vi specifiera storleken på p.

Definitionerna (8.5), (8.6) och (8.7) uppfyller sambandet (8.4) eftersom

$$\frac{N_k + \delta_k}{N} = \frac{\frac{1}{(p-1)!} \int_0^\infty e^{k-x} f(x) dx}{\frac{1}{(p-1)!} \int_0^\infty e^{-x} f(x) dx} = e^k.$$

Låt oss nu definiera talen $c_0, c_1, c_2, \ldots, c_{np} \in \mathbb{Z}$, sådana att

$$\begin{aligned}((x-1)(x-2)\ldots(x-n))^p &= (x^n + \cdots \pm n!)^p = x^{np} + \cdots \pm (n!)^p \\ &= \sum_{j=0}^{np} c_j x^j\end{aligned}$$

med $c_{np} = 1$ och $c_0 = \pm(n!)^p$. Notera att $\pm(n!)^p$ är inte delbar med p eftersom $p > n$.

Vi skriver om N, N_k och δ_k turvis genom att använda $f(x) = x^{p-1} \sum_{j=0}^{np} c_j x^j$. Vi börjar med N från (8.5)

$$\begin{aligned}N &= \frac{1}{(p-1)!} \int_0^\infty e^{-x} f(x) dx \\ &= \frac{1}{(p-1)!} \sum_{j=0}^{np} c_j \int_0^\infty x^{j+p-1} e^{-x} dx.\end{aligned}$$

Integralen $\int_0^\infty x^{j+p-1} e^{-x} dx$ kan skrivas som $(j+p-1)!$ genom att använda gammafunktionen (8.1). Vi får

$$\begin{aligned}N &= \sum_{j=0}^{np} c_j \frac{(j+p-1)!}{(p-1)!} \\ &= c_0 \frac{(0+p-1)!}{(p-1)!} + \sum_{j=1}^{np} c_j \frac{(j+p-1)!}{(p-1)!} \\ &= c_0 + \sum_{j=1}^{np} c_j \frac{(j+p-1)!}{(p-1)!}.\end{aligned}$$

Summan $\sum_{j=1}^{np} c_j \frac{(j+p-1)!}{(p-1)!}$ är en multipel av p eftersom den innehåller termer bestående av p, men som tidigare nämnt är c_o inte delbar med p eftersom $p > n$. Detta medför att N inte heller är delbar med p. Därmed $p \nmid N$.

Vi fortsätter med N_k. Substitution av $t = x - k$ i (8.6) ger

$$\begin{aligned} N_k &= \frac{1}{(p-1)!} \int_0^\infty e^{-t} f(t+k) dt \\ &= \frac{1}{(p-1)!} \sum_{j=1}^{np} c_j \int_0^\infty (t+k)^{j+p-1} e^{-t} dt \\ &= \sum_{j=1}^{np} c_j \frac{(j+p-1)!}{(p-1)!}. \end{aligned}$$

Summan $\displaystyle\sum_{j=1}^{np} c_j \frac{(j+p-1)!}{(p-1)!}$ är en multipel av p av samma anledning som nämnt ovan, alltså delar p denna summa. Med andra ord är N_k delbar med p, uttryckt som $p \mid N_k$.

Låt oss gå vidare med de sista talen δ_k i (8.7) och undersöka deras storlek. Vi har att

$$\begin{aligned} |\delta_k| &= \frac{1}{(p-1)!} \left| \int_0^k e^{k-x} f(x) dx \right| \\ &\leq \frac{e^k}{(p-1)!} \int_0^k |f(x)| \, e^{-x} dx. \end{aligned}$$

Absolutbeloppet av funktionen $f(x)$ är en kontinuerlig funktion som är begränsad i $x \in [0, n]$, vilket innebär att $|f(x)| \leq A$ för något heltal A. Vi får

$$\begin{aligned} |\delta_k| &\leq \frac{e^k A}{(p-1)!} \int_0^k e^{-x} dx \\ &\leq \frac{e^k A}{(p-1)!} \int_0^\infty e^{-x} dx \\ &= \frac{e^k A}{(p-1)!}. \end{aligned}$$

Vi har en exponentialfunktion $e^k A$ i täljaren och en fakultet i nämnaren, vilket medför att uttrycket δ_k går mot noll när vi väljer allt större primtal p och låter det gå mot oändlighet. Alltså $\lim_{p \to \infty} \delta_k = 0$.

Vi återknyter till polynomekvationen (8.3) och sätter in sambandet kring e^k (8.4).

$$\begin{aligned} & a_n e^n + \cdots + a_1 e + a_0 = 0 \\ & a_n \left(\frac{N_n + \delta_n}{N} \right) + \cdots + a_1 \left(\frac{N_1 + \delta_1}{N} \right) + a_0 = 0 \\ & a_n \left(N_n + \delta_n \right) + \cdots + a_1 (N_1 + \delta_1) + a_0 N = 0 \\ & (a_n N_n + \cdots + a_1 N_1 + a_0 N) + (a_n \delta_n + \cdots + a_1 \delta_1) = 0 \end{aligned} \tag{8.8}$$

Uttrycket $(a_nN_n + \cdots + a_1N_1 + a_0N)$ är inte multipel av p eftersom p inte delar a_0N. Eftersom talet 0 är en multipel av p innebär detta att uttrycket i fråga är inte lika med 0, men det är ett heltal då dess samtliga termer är heltal. Därutöver går uttrycket $(a_n\delta_n + \cdots + a_1\delta_1)$ mot noll för att termen δ_k för $k = 1, 2, \ldots, n$ som tidigare nämnt närmar sig noll när p går mot oändligheten. Eftersom mängden av primtal är oändlig kan vi välja ett godtyckligt stort primtal p som medför att värdet av uttrycket $(a_n\delta_n + \cdots + a_1\delta_1)$ hamnar mellan 0 och 1. För att likheten (8.8) ska gälla måste dock båda uttryck vara lika med eller närma sig noll, men i själva verket är uttrycket $(a_nN_n + \cdots + a_1N_1 + a_0N) \neq 0$. Vi har därmed funnit en motsägelse, vilket innebär att e uppfyller ingen polynomekvation med icke-noll heltalskoefficienter. Med andra ord är e ett transcendent tal. $\square$

9 Approximation av e

I detta avsnitt ska vi approximera e genom gränsvärde, serie och kedjebråk med hjälp av SageMath [17]. Vi har valt att använda SageMath eftersom den kan hantera flyttal med godtycklig precision, och har koder för kedjebråk och konvergenter som är redan definierade utan att importera något bibliotek. Vi ska först jämföra gränsvärde med serie och därefter serie med kedjebråk.

9.1 Reella approximationer

Vi börjar att approximera e med gränsvärdet (2.4), alltså

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n.$$

Resultatet presenteras i form av en tabell där ett antal värden på n har valts ut och approximationsfelet för respektive värde har beräknats genom att ta absolutbeloppet av $\Delta = \left(1+\frac{1}{n}\right)^n - e$ för de olika värdena på n.

n	**Approximationsfel** $\lvert\Delta\rvert$
0	1,7182818
1	0,7182818
10	0,12453937
100	0,013467999
1000	0,0013578962
10^5	1,3591284555567326e-05
10^6	1,3591396683515404e-06
10^{12}	1,3591409142282767e-12

Låt oss ta reda på hur stort värdet på n bör vara för att få 100 korrekta decimaler. För att beräkna detta behöver vi tillämpa Maclaurinutvecklingarna för funktionerna $\ln(1+x)$ och e^x, vilka är som följande

$$\ln(1+x) = x - \frac{x^2}{2} + O(x^2), \tag{9.1}$$
$$e^x = 1 + x + O(x^2), \tag{9.2}$$

där $O(x^2)$ är resttermerna. Vi inför beteckningen $x = 1/n$ och får att

$$\left(1+\frac{1}{n}\right)^n = e^{\ln(1+x)^{\frac{1}{x}}} = e^{\frac{1}{x}\ln(1+x)}.$$

Substitutera in Maclaurinutvecklingen för $\ln(1+x)$ från (9.1) vilket ger

$$\begin{aligned}\left(1+\frac{1}{n}\right)^n &= e^{\frac{1}{x}(x-\frac{x^2}{2}+O(x^2))}\\ &= e^{1-\frac{x}{2}+O(x^2)}\\ &= e\cdot e^{-\frac{x}{2}+O(x^2)}.\end{aligned}$$

Substitutera in även Maclaurinutvecklingen för e^x genom att byta ut x mot $-x/2+O(x^2)$ så att vi får

$$\begin{aligned}\left(1+\frac{1}{n}\right)^n &= e\cdot\left(1-\frac{x}{2}+O(x^2)\right)\\ &= e\cdot\left(1-\frac{1}{2n}+O\left(\frac{1}{n^2}\right)\right)\\ &= e-\frac{e}{2n}+O\left(\frac{1}{n^2}\right).\end{aligned}$$

För att beräkna hur stor n bör vara för att vi ska få 100 korrekta decimaler använder vi formeln $|\Delta| < 0,5\cdot 10^{-t}$ där Δ är skillnaden av närmevärdet och det korrekta värdet medan t står för antal korrekta decimaler [18]. Så ifall vi vill ha 100 korrekta decimaler sätter vi $t = 100$, därmed

$$\left|\left(1+\frac{1}{n}\right)^n - e\right| = \left|e-\frac{e}{2n}+O\left(\frac{1}{n^2}\right)-e\right| = \frac{e}{2n}+O\left(\frac{1}{n^2}\right) < 0,5\cdot 10^{-100}.$$

Vi bortser från resttermen eftersom den blir mycket liten för stora n och tittar närmare på olikheten

$$\frac{e}{2n} < \frac{1}{2\cdot 10^{100}} \Rightarrow n > e\cdot 10^{100}.$$

Alltså måste n vara större än eller lika med ungefär $2,8\cdot 10^{100}$ för att vi ska kunna få en approximation av e med 100 korrekta decimaler.

Låt oss nu approximera e med serien (2.5), alltså

$$\sum_{k=0}^{\infty}\frac{1}{k!}.$$

Vi utvecklar serien och får att

$$e = \sum_{k=0}^{\infty}\frac{1}{k!} = \sum_{k=0}^{n}\frac{1}{k!} + \sum_{k=n+1}^{\infty}\frac{1}{k!}$$

där summan $R_n = \sum_{k=n+1}^{\infty} \frac{1}{k!}$ är resten, alltså är det approximationsfelet när vi tar absolutbeloppet av skillnaden mellan delsumman $\sum_{k=0}^{n} \frac{1}{k!}$ och e för olika värden på n. Låt oss alltså fokusera på resten för att approximera felet. Vi har

$$R_n = \sum_{k=n+1}^{\infty} \frac{1}{k!} = \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \cdots = \frac{1}{n!}\left(\frac{1}{n+1} + \frac{1}{(n+1)(n+2)} + \dots\right).$$

Uttrycket $\frac{1}{n+1} + \frac{1}{(n+1)(n+2)} + \dots$ är ekvivalent med uttrycket från (5.4) som visade sig vara mindre än $1/b$. Ifall vi byter b mot n får vi alltså

$$R_n = \frac{1}{n!}\left(\frac{1}{n+1} + \frac{1}{(n+1)(n+2)} + \dots\right) \leq \frac{1}{n!} \cdot \frac{1}{n}. \tag{9.3}$$

Alltså är approximationsfelet $|\Delta| = \left|\sum_{k=0}^{n} \frac{1}{k!} - e\right| \leq \frac{1}{n!} \cdot \frac{1}{n}$.

För att beräkna hur stor n måste vara för att vi ska få 100 korrekta decimaler använder vi SageMath och får att $n = 70$ ger ett approximationsfel på

$$|\Delta| = 1,1923660008307707 \cdot 10^{-102} < 0,5 \cdot 10^{-100}.$$

Även när vi använder uttrycket för approximationsfelet från (9.3) får vi ungefär samma resultat, alltså att $n = 70$ ger

$$\frac{1}{70!} \cdot \frac{1}{70} = 1,1926058197346155 \cdot 10^{-102} < 0,5 \cdot 10^{-100}.$$

I jämförelse med gränsvärde där vi behöver ha $n \approx 2,8 \cdot 10^{100}$ för att få en approximation av e med 100 korrekta decimaler kan vi alltså med hjälp av serien (2.5) uppnå samma resultat med endast $n = 70$. Med andra ord konvergerar serien (2.5) mycket snabbare mot e än gränsvärdet (2.4).

9.2 Rationella approximationer

Vi kommer nu approximera e med kedjebråksutvecklingen (5.5) genom att välja ut index (n) för olika konvergenter $c_n = p_n/q_n$ och uppskatta deras approximationsfel $|\Delta| = |p_n/q_n - e|$. I synnerhet ska vi fokusera på konvergenternas nämnare q_n för diverse n.

Vi börjar med att approximera felet. Vi minns från avsnittet om kedjebråk att konvergenter med udda index c_{n+1} är större, medan de med jämn index c_n är mindre än ett tal a. Alltså ligger talet mellan dessa två konvergenter. Det gäller alltså att $|c_n - a| \leq |c_n - c_{n+1}|$. Vi vet från (4.12) att $c_{n+1} - c_n = \dfrac{(-1)^n}{q_n q_{n+1}}$ vilket ger att

$$|c_n - a| \leq |c_n - c_{n+1}| = \frac{1}{q_n q_{n+1}}.$$

Eftersom q_n från 4.10 definieras av en rekursiv talföljd vars termer är positiva innebär detta att $q_n < q_{n+1}$. Alltså får vi

$$|c_n - a| < \frac{1}{{q_n}^2}.$$

Med andra ord är approximationsfelet mindre än kvoten $1/{q_n}^2$. I följande tabell presenterar vi approximationsfelet $|\Delta|$, samt produkten av $|\Delta|$ och nämnaren q_n för konvergenter av olika index n

n	**Nämnare** q_n	**Approximationsfel** $\|\Delta\|$	$\|\Delta\| \cdot {q_n}^2$
0	1	0,7182818284590452	0,7182818284590452
1	1	0,28171817154095474	0,28171817154095474
5	32	0,0004681715409547646	0,479407657937679
10	1001	1,1017732695364201e-07	0,11039779178487624
15	208524	1,153848641754146e-11	0,5017194499832017
20	150869313	2,18383037131314e-17	0,4970736332575045
25	10622799089	4,657588358181949e-22	0,05255802508884514
30	8690849042711	6,626631462804679e-27	0,5005151539600752

Vi kan påminna oss om satsen 4.17 från avsnittet om kedjebråk, där det framgår att ifall ett rationellt tal p/q har ett approximationsfel $|\Delta| = |\alpha - p/q|$ som är mindre än $1/2q^2$ så är den en konvergent i kedjebråksutvecklingen av det irrationella talet α. I den meningen ser vi även i tabellen att produkten av approximationsfelet $|\Delta|$ och q^2 är mindre än $0,5$ för vissa konvergenter. Detta innebär att de uppfyller villkoret $|\Delta| \cdot 2q^2 < 1$ som finns i satsen 4.17. Alltså är approximationsfelet minimalt. Konvergenter i kedjebråksutvecklingen av ett irrationellt tal är de bästa rationella approximationerna av talet eftersom de ger en approximation som är omöjligt att uppnå med ett annat rationellt tal med mindre nämnare.

På liknande sätt som tidigare vill vi få en approximation av e med 100 korrekta decimaler. I detta fall ska vi därmed finna den n:e konvergent som ger en sådan approximation med hjälp av SageMath. Vi får att $n = 88$, alltså att den 88:e konvergenten ger ett approximationsfel på

$$|\Delta| = 4,54691513308561 \cdot 10^{-101} < 0,5 \cdot 10^{-100}.$$

Med andra ord får vi 100 korrekta decimaler med den 88:e konvergenten.

Låt oss jämföra effektiviteten av de två rationella approximationerna kedjebråk och serie. Om vi jämför storleken av kedjebråkets respektive seriens nämnare och täljare för ett antal korrekta decimaler kan vi notera en stor skillnad. För att illustrera detta kan vi exempelvis ta reda på den rationella approximationen som ger 27 korrekta decimaler med kedjebråk och serie. Vi får att $\frac{10962598503531495302220342777}{4032914611266056355840000000}$ och $\frac{46150226651233}{16977719590391}$ ger 27 korrekta decimaler med serie respektive kedjebråk. Notera att vi får en mycket större täljare och nämnare med serie i jämförelse med kedjebråk. Exempelvis finns det 28 siffror i täljaren av det första bråket medan det är hälften så många i det andra bråket. Detta beror på att konvergenten $c_n = p_n/q_n$ i kedjebråksutvecklingen av e har ett fel av storleksordning $1/{q_n}^2$, medan delsumman $\sum_{k=0}^{n} \frac{1}{k!} = \frac{a_n}{b_n}$ som vi strax kommer att se har ett fel som är runt $1/b_n$.

Låt oss presentera produkten av nämnare b_n med ett urval index n och respektive approximationsfel i form av en tabell.

n	**Nämnare b_n**	$\lvert\Delta\rvert \cdot b_n$
0	1	1,7182818284590453
1	1	0,7182818284590452
5	60	0,09690970754271412
10	3628800	0,09911218335007542
15	326918592000	0,01659800829379285
20	2432902008176640000	0,04988174288517625
25	3139921061403033600000	8,084763091747005e-05
30	204040661393993122027929600000000	0,0025613539460002457

Från tabellen kan vi uppmärksamma att felet är mycket större än $1/{b_n}^2$. Det verkar snarare vara av storleksordning $1/b_n$, vilket skulle kunna förklara varför vi får dubbelt så stora täljare och nämnare med serien jämfört med kedjebråket.

För att sammanfatta är konvergenter i kedjebråksutvecklingen av ett tal de bästa rationella approximationerna. Som vi kan se kan man inte få en bättre approximation med en annan rationell approximation som har mindre nämnare och täljare. Exempelvis ger konvergenten 27 decimalers noggrannhet med hälften så många tal som delsumman. Alltså är konvergenten en bättre rationell approximation än delsumman. Därutöver kan vi även jämföra konvergensen av serie respektive kedjebråk. Vi fick att $n = 70$ ger 100 korrekta decimaler med serien medan den 88:e konvergenten ger samma noggrannhet. Alltså konvergerar serien mycket snabbare mot e än kedjebråket, vilket kan förklaras av den snabbväxande nämnaren $n!$.

10 Avslutande ord

Syftet med arbetet var att fördjupa förståelsen av talet e. I detta arbete har vi därmed utforskat talets historia, bevisat att den är ett irrationellt och till och med ett transcendent tal, härlett dess kedjebråk, samt redogjort för och jämfört några av dess approximationsmetoder. Jag hoppas att det har varit givande att läsa och medfört ökade kunskaper om talet e.

Referenser

[1] Wikipedia (2024). Natural logarithm. Tillgänglig här

[2] O'Connor, J.J., Robertson, E.F. (2001). The number e. Tillgänglig här

[3] Maor, E. (1994). *e: The Story of a Number.* Princeton: Princeton University Press. Tillgänglig här

[4] Thunberg, H. (u.å.). Kompletterande kurslitteratur om serier. Kungliga Tekniska Högskolan. Tillgänglig här

[5] Monks, K.M. (2022). Why $\sqrt{2}$ is a Friendlier Number than e: Irrational Adventures with Aristotle, Fourier, and Liouville. *TRIUMPHS Digital Commons Collection. Calculus.* 22. Tillgänglig här

[6] LetsSolveMathProblems (2018). Fourier's Proof that e is Irrational. Youtube. Tillgänglig här

[7] Wyman, M.F., Wyman, B.F. (1985). An Essay on Continued Fractions. *Mathematical Systems Theory.* 18. Tillgänglig här

[8] Tambour, T. (2002). Kedjebråk: nästan bortglömd matematik. Stockholms universitet, Matematiska institutionen.

[9] Olds, C.D. (1963). *Continued fractions.* Washington: The Mathematical Association of America. Anneli Lax New Mathematical Library, vol. 9. Tillgänglig här

[10] Persson, A., Böiers, L.C. (2010). *Analys i en variabel.* 3:e upplaga. Lund: Studentlitteratur.

[11] Sandifer, E. (2006). Who proved e is irrational? *Mathematical Association of America.* Tillgänglig här

[12] Cohn, H. (2006). A Short Proof of the Simple Continued Fraction Expansion of e. *American Mathematical Monthly,* vol. 113. Tillgänglig här

[13] Ross, M. (2018). e and π are transcendental. Tillgänglig här

[14] Penn, M. [MichealPennMath] (2022). e is transcendental – the best proof! Youtube. Tillgänglig här

[15] Polster, B. [Mathologer] (2018). The PROOF: e and pi are transcendental. Youtube. Tillgänglig här

[16] Wikipedia (2024). Gammafunktionen. Tillgänglig här

[17] SageMath, the Sage Mathematics Software System (Version 10.4), The Sage Developers, 2024, www.sagemath.org

[18] Wikipedia (2017). Korrekta decimaler. Tillgänglig här