UNIVERSIDADE FEDERAL DO CEARÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
CENTRO DE CIÊNCIAS
MESTRADO PROFISSIONAL EM MATEMÁTICA

LANA PRISCILA SOUZA

CRIPTOGRAFIA RSA: A TEORIA DOS NÚMEROS POSTA EM PRÁTICA

FORTALEZA
2015

LANA PRISCILA SOUZA

CRIPTOGRAFIA RSA: A TEORIA DOS NÚMEROS POSTA EM PRÁTICA

Dissertação de Mestrado apresentada ao Programa de Mestrado Profissional em Matemática da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Matemática. Área de concentração: Matemática.

Orientador: Prof. Dr. José Othon Dantas Lopes

FORTALEZA

2015

Página reservada para ficha catalográfica que deve ser confeccionada após apresentação e alterações sugeridas pela banca examinadora.

Para solicitar a ficha catalográfica de seu trabalho, acesse o site: www.biblioteca.ufc.br, clique no banner Catalogação na Publicação (Solicitação de ficha catalográfica)

LANA PRISCILA SOUZA

CRIPTOGRAFIA RSA:
A TEORIA DOS NÚMEROS POSTA EM PRÁTICA

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Matemática em Rede Nacional, do Departamento de Matemática da Universidade Federal do Ceará, como requisito parcial para a obtenção do Título de Mestre em Matemática. Área de concentração: Ensino de Matemática.

Aprovada em: 11 / 06 / 2015.

BANCA EXAMINADORA

Prof. Dr. José Othon Dantas Lopes (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. José Robério Rogério
Universidade Federal do Ceará (UFC)

Prof. Dr. Ângelo Papa Neto
Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

Dedico este trabalho à minha família e aos meus amigos pelo apoio de sempre em todos os meus projetos.

AGRADECIMENTOS

A Deus, por cada dia.

À minha mãe Margarida, por acreditar em mais esse projeto. Grata pela compreensão, calma e fé com que sempre me ensinou a levar a vida!

Aos meus familiares, em especial aos meus irmãos Patrícia e Diego (a de verdade e o de coração), que dão força e torcem por minhas conquistas. Agradeço por sua generosidade em todos os momentos!

À CAPES, pelo apoio financeiro com a manutenção da bolsa de auxílio.

Ao Prof. Dr. José Othon Dantas Lopes, pelo caminho percorrido desde a escolha do tema. Grata pela disponibilidade e serenidade com que conduziu o trabalho.

Aos professores Robério e Papa, que leram meu trabalho, participaram da banca e, generosamente, me presentearam com dicas valorosas.

Aos professores e colegas da turma de mestrado, principalmente ao meu amigo Uchoa, pela presença, por tentar me dar força e tranquilizar nos momentos complicados do curso, e por sua enorme disponibilidade em ajudar sempre.

À amiga Cristiane, que compreendeu todas as minhas angústias (aquelas que escutou e todas as outras das quais eu não disse uma palavra).

A todos os amigos que mais uma vez acompanharam todo o processo de um trabalho árduo e demorado.

“ – É sobre criptografia...

– Como mensagens secretas?

– Não secretas. Essa é a parte brilhante. Mensagens que todos podem ver, mas ninguém sabe o que são, a menos que tenha a chave.

– Como isso é diferente de falar?

– Falar?

– Quando as pessoas conversam, nunca dizem o que querem. Dizem outra coisa... E esperam que você saiba o que querem dizer. Só que eu nunca sei. Então, como isso é diferente? ”

(JOGO..., 2014)

RESUMO

Desde o advento da escrita, o envio de mensagens secretas tem sido uma importante maneira de guardar sigilo de informações confidenciais. A arte de elaborar mensagens a partir de códigos secretos surge na figura da criptografia que, com o passar do tempo, estende os seus serviços às transações comerciais realizadas pela *internet*. O principal algoritmo utilizado pela *internet* recebe o nome de RSA. Assim, a criptografia RSA codifica números de cartões de créditos, senhas de bancos, números de contas e utiliza para isso elementos de uma importante área da Matemática: a Teoria dos Números.

Palavras-chave: Primos. Criptografia. RSA.

ABSTRACT

Since the advent of writing, sending secret messages has been an important way to maintain confidentiality of sensitive information. The art of crafting messages from secret codes appears in the figure of encryption that over time extends its services to commercial transactions over the Internet. The main algorithm used by the internet is called RSA. Thus, the RSA Encryption encodes credit card numbers, bank passwords, account numbers and uses for that elements of an important area of mathematics: number theory.

Keywords: Prime. Encryption. RSA.

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

1 INTRODUÇÃO

Possuindo apenas dois divisores, os números primos têm a importante propriedade de gerar todos os outros números e, por isso, são considerados os átomos da aritmética. A história é cercada de peças que parecem pertencer e/ou ajudar a montar o gigante e incompleto quebra-cabeça dos primos. Diversos matemáticos marcaram seu nome nessa história, seja estudando a primalidade dos números, provando ideias deixadas por outros matemáticos ou tentando encontrar uma fórmula para ordená-los. A esse respeito, Sautoy (2007, p. 14) destaca:

> a questão tem atormentado as mentes matemáticas de todas as épocas. Depois de mais de dois mil anos de esforços, os primos parecem resistir a qualquer tentativa de encaixá-los em um padrão reconhecível. O tambor dos primos tem tocado sua sequência de números ao longo de gerações: duas batidas, seguidas por três batidas, cinco, sete, onze. O ritmo segue em frente, e torna-se fácil acreditar que seja causado por ruído branco aleatório, sem qualquer lógica interna. No centro da matemática, que é a busca pela ordem, só escutávamos o som do caos.

Fora a ordenação dada por impossível, nos deparamos com outra questão que atormenta as mentes dos estudiosos de todos os tempos: a fatoração em primos. Já que os primos atuam como blocos para a construção de todos os outros números, como determinar os primos que compõem determinado número composto? Não que toda decomposição em primos seja impossível, longe disso! Aprendemos a utilizar fatoração de primos na escola para determinar, entre outras coisas, o máximo divisor comum (MDC) e o mínimo múltiplo comum (MMC), mas nossos trabalhos restringem-se a números didáticos e de simples compreensão.

A grande questão que envolve a decomposição ou fatoração em primos diz respeito ao trabalho com números extensos (números com uma quantidade grande de algarismos). Não se tem notícia até hoje de um algoritmo que apresente baixo custo computacional e que permita a fatoração em primos de números extensos e isso, em plena era da informática, onde encontramos computadores tão potentes, é, no mínimo, questionável. Pensando nisso, Lovász, Pelikán e Vesztergombi (2013, p. 92) colocam:

> é claro que supercomputadores poderosos e sistemas massivamente paralelos podem ser usados para encontrar decomposições por meio da força bruta para números um tanto grandes; o recorde atual é cerca de 140 dígitos, e a dificuldade cresce muito rapidamente (exponencialmente) com o número de dígitos. Encontrar a decomposição prima de um dado número com 400 dígitos, por qualquer dos

> métodos conhecidos, está muito além das possibilidades dos computadores no futuro previsível.

É importante destacar que o problema do custo computacional que impossibilita a fatoração de números extensos (como esses que o autor menciona) não deixou os estudiosos parados, ao contrário. A busca por um algoritmo com baixo custo que permita tal fatoração continua. E enquanto tal algoritmo não é encontrado, os números extensos difíceis de ser fatorados atendem às necessidades da codificação de informações importantes (como números de catão, senhas, contas bancárias, etc.) enviadas pela *internet*.

Assim, graças a uma importante descoberta feita por um trio de integrantes do MIT (Instituto de Tecnologia de Massachusetts), a utilização dos primos e o "problema da fatoração de números extensos" apresentam uma aplicação prática que se estendeu à *internet* – a criptografia RSA. O nascimento da criptografia na *internet* marca um período de avanços notórios na fase emergente da célere comunicação global.

Desse modo, o trabalho objetiva a compreensão da criptografia RSA que tem por base a escrita e envio de mensagens por meio de um algoritmo que faz uso dos números primos e de outros elementos fundamentais da Teoria dos Números. Nesse contexto, o trabalho consta da apresentação do histórico da criptografia até o desenvolvimento da RSA, ponto onde a matemática passa a ter um papel de grande destaque.

Com caráter bibliográfico, o trabalho procura basear-se em diversos autores como Sautoy, Hefez e Coutinho, que tratam do assunto e procuram apresentar perspectivas, visões críticas e exemplos concretos que, quando combinados, geram possibilidades para o uso da criptografia RSA. De acordo com Gil (2008, p. 50):

> a pesquisa bibliográfica é desenvolvida a partir de material já elaborado, constituído principalmente de livros e artigos científicos. [...] Parte dos estudos exploratórios podem ser definidos como pesquisas bibliográficas, assim como certo número de pesquisas desenvolvidas a partir da técnica de análise de conteúdo. A principal vantagem da pesquisa bibliográfica reside no fato de permitir ao investigador a cobertura de uma gama de fenômenos muito mais ampla do que aquela que poderia pesquisar diretamente.

Além disso, a pesquisa bibliográfica "também é indispensável nos estudos históricos. Em muitas situações, não há outra maneira de conhecer os fatos passados senão com base em dados secundários" (GIL, 2008, p. 51). Consequentemente, a pesquisa procura realizar uma revisão bibliográfica objetivando situar o leitor no histórico dos números primos

e da criptografia, apresentando o algoritmo que utiliza alguns exemplos práticos e motivos que garantem seu funcionamento e segurança.

Dividido em quatro capítulos, o trabalho trata da criptografia RSA como uma prática da Teoria dos Números. No primeiro capítulo, apresentaremos uma breve exposição sobre o histórico dos números primos; o segundo capítulo trata de alguns conceitos da Teoria dos Números; o terceiro, descreve o percurso da criptografia ao longo do tempo, além de apresentar o desenvolvimento da criptografia RSA; o último capítulo trata da parte prática da RSA, apresentando o algoritmo que permite a codificação e decodificação utilizadas principalmente nas transações comerciais realizadas pela *internet*.

Assim, nos propomos a expor pontos relevantes à compreensão da criptografia RSA, levando em conta todo o processo de codificação e decodificação de mensagens, enfatizando a presença indispensável da matemática para seu desenvolvimento e aplicabilidade.

2 CAPÍTULO I – NÚMEROS PRIMOS: DEFINIÇÃO, HISTÓRICO E O TEOREMA FUNDAMENTAL DA ARITMÉTICA

Os números primos vêm se mostrando desde sempre como um mistério para o mundo matemático. Como podem números definidos de forma tão simples revelar problemas tão complexos? A história desses números é cercada de quebra-cabeças, aparentemente indissolúveis, muitos dos quais só foram provados séculos depois de serem descobertos.

Nesse sentido, o capítulo apresenta o conceito de números primos e uma perspectiva histórica de seu estudo, além de destacar um importante teorema que exibe os números primos como os elementos básicos à construção de todos os outros números – o teorema fundamental da aritmética.

2.1 Definição

Daremos início ao tópico definindo o nosso principal objeto de estudo: os números primos. Um inteiro positivo $p > 1$ é primo se seus únicos divisores positivos forem 1 e p. Um inteiro $p > 1$ que não é primo é dito composto. Assim, por exemplo, os números 2, 3, 5, 7, 11 e 13 são primos e os números 6, 8, 10, 15 e 20 são compostos. O número 1 não é considerado nem primo e nem composto.

Apesar da simplicidade e naturalidade enunciadas na definição, o conjunto dos números primos surge envolto de mistérios e enigmas que, apesar de terem premissas deveras despretensiosas, são capazes de dar um trabalho de extensas proporções ao matemático mais proeminente. Questões de simples elaboração e ideias de aparência elementar permanecem incógnitas até hoje. É claro que a busca por desvendar esses mistérios continua, mesmo porque:

> os matemáticos não suportam admitir a possibilidade de que talvez não exista uma explicação para o modo como a natureza escolheu os primos. Se a matemática não tivesse uma estrutura, uma simplicidade bela, não valeria a pena estudá-la. Escutar ruído branco nunca foi um passatempo muito apreciado. [...] Os números primos representam para os matemáticos um dos dilemas mais estranhos de sua disciplina. (SAUTOY, 2007, p. 14-15)

Desse modo, uma vez definido o principal ponto do trabalho, trataremos a seguir de uma breve trajetória sobre esses números tão instigantes, apontando alguns matemáticos que deram importantes contribuições à busca histórica por números primos.

2.2 Histórico

Tendo sua história tão antiga quanto a história da descoberta dos números, o conjunto denominado de números primos apresenta-se como um dos objetos mais misteriosos de que se tem notícia no mundo matemático.

Diversos matemáticos ocuparam-se com a observação e a descoberta de propriedades sobre o conjunto dos primos. Os antigos gregos já desenvolviam estudos nessa área, que mais tarde ficou conhecida como Teoria dos Números, mas, de acordo com Sautoy (2007, p. 31):

> algumas pessoas acreditam que os chineses tenham sido a primeira cultura a escutar o ritmo do tambor dos primos. Eles atribuíam características femininas aos números pares e masculinas aos ímpares. Além dessa divisão conservadora, também existiam os números afeminados, formados pelos números ímpares que não são primos, como 15. Há indícios de que em 1000 a.C. os chineses já haviam desenvolvido um método bastante físico de entender o que torna os primos tão especiais em relação a todos os demais números. [...] Para os chineses, os primos eram os números machões que resistiam a qualquer tentativa de separação em um conjunto de números menores.

Foram os gregos que, já suspeitando da infinidade dos primos, descobriram que eles podiam gerar todos os outros números. Uma descoberta deveras fascinante já que o conjunto dos primos se mostra, por vezes, impenetrável. Sautoy (2007, p. 31), a esse respeito, destaca:

> os gregos da antiguidade também gostavam de atribuir qualidades sexuais aos números, mas foram eles que descobriram, no quarto século a.C., a capacidade dos primos de servir como blocos de construção para todos os números. Eles perceberam que todo número podia ser gerado pela multiplicação de números primos.

A descoberta grega abriu as portas para a compreensão de vários pontos estudados posteriormente na Teoria dos Números, mas não foi o bastante para que pudéssemos identificar a chamada primalidade dos números. Sautoy (2007, p.31) completa: “Apesar do rápido sucesso dos gregos na identificação dos blocos de construção da aritmética, os matemáticos ainda têm dificuldades para entender a tabela de números primos”.

Acredita-se que várias tentativas foram desenvolvidas com o objetivo de identificar números primos e de criar uma tabela para representá-los. Eratóstenes (Figura 1) é um dos matemáticos mais importantes nesse estudo e a quem deve-se destaque.

Figura 1

Fonte: Disponível em: <http://www.sandysnunes.com.br/2013/09/gerando-numeros-primos-com-scala-crivo.html>. Acesso em: 25/02/2015.

Nascido em Cirene, na costa sul do Mediterrâneo, Eratóstenes era um pouco mais novo que Arquimedes. O matemático passou grande parte da vida em Atenas, mas foi convidado por Ptolomeu III do Egito, quando tinha cerca de 40 anos, a mudar-se para Alexandria com o objetivo de ser tutor de seu filho e bibliotecário chefe da universidade local.

Eves (2004, p. 197) relata que Eratóstenes foi singularmente talentoso e destacou-se como matemático, astrônomo, historiador, geógrafo, filósofo, poeta e atleta. Conta ainda que, por volta de 194 a.C., já com idade avançada, foi acometido por uma oftalmia que o deixou cego. O matemático, muito abalado com a cegueira, resolveu suicidar-se, deixando voluntariamente de se alimentar.

Tendo dedicado parte de sua vida à matemática, Eratóstenes nos deixou valorosas contribuições. Nos relatos de Sautoy (2007, p. 31) fica evidenciado que:

> até onde sabemos, a primeira pessoa a produzir tabelas de números primos foi o diretor da biblioteca do grande instituto de pesquisa da Grécia Antiga, localizada em Alexandria. Como um Mendeleiev matemático ancestral, Eratóstenes descobriu, no terceiro século a.C., um procedimento relativamente indolor para determinar quais números são primos em uma lista que inclua, por exemplo, os primeiros mil números. Eratóstenes escrevia inicialmente uma lista com todos os números de 1 a 1000. Em seguida, escolhia o primeiro primo, 2, e eliminava da lista todos os seus múltiplos. Como todos esses números eram divisíveis por 2, obviamente não eram primos. Logo, passava ao seguinte número que não fora eliminado, ou seja, o número 3, e eliminava também todos os seus múltiplos. Como todos eram divisíveis por 3, tampouco eram primos. Eratóstenes foi em frente, escolhendo sempre o seguinte número que não havia sido retirado da lista e eliminando todos os números divisíveis por esse novo primo. Com esse processo sistemático ele produziu tabelas de primos. Mais tarde, o procedimento passou a ser chamado de *crivo de Eratóstenes*. Cada novo primo gerava um "crivo" que Eratóstenes utilizava para eliminar os números não primos. O tamanho do crivo se alterava em cada etapa, mas ao atingir o número 1000, somente os números primos resistiam a todos os crivos.

O crivo que Eratóstenes nos deixou lista os números primos menores ou iguais a um natural *n*. Além disso, Eratóstenes também forneceu um método para determinar se um dado número *n* é primo. Para tanto, basta-nos verificar que o número em questão não é divisível por nenhum primo *p* que não supere $\sqrt{n}$. A proposição a seguir mostra-nos tal fato.

Proposição 2.1.1 Se *n* é composto, o seu menor divisor diferente de 1 não é maior que $\sqrt{n}$.

Demonstração. Queremos mostrar que se *n* não possuir divisores diferentes de 1, menores ou iguais a $\sqrt{n}$, então *n* é primo. Para isso, considere *p* como o menor divisor de *n* diferente de 1. Portanto, $n = pq$ com $q \geq p$. Multiplicando ambos os membros da desigualdade por *p*, obtemos $n = pq \geq p^2$ e, consequentemente, $\sqrt{n} \geq p$.

Apesar disso, somos levados a repetir que mesmo que a proposição tenha auxiliado no trabalho com a determinação de um primo, até ela tem um limite e esse limite é dado pela quantidade de dígitos de um dado número. Para uma grande quantidade de dígitos, a busca pela primalidade, por meio do crivo e da proposição, mostra-se impraticável.

Mesmo assim, a descoberta de números primos tornou-se o passatempo preferido de muitos matemáticos e curiosos. Seja programando o crivo de Eratóstenes em computador, seja utilizando outros métodos, a descoberta de primos parece, apesar de exaustiva, verdadeiramente excitante. Euclides de Alexandria (Figura 2) também deu sua contribuição na busca incansável pelos números primos.

Figura 2

Fonte: Disponível em: < http://www.geometriaanalitica.com.br/artigos/euclides.html>. Acesso em: 25/02/2015.

Mestre, escritor de origem provavelmente grega, matemático da escola platônica e conhecido como o Pai da Geometria, Euclides nasceu na Síria e realizou seus estudos em Atenas. É até hoje, na história da Matemática, considerado como um dos mais significativos estudiosos deste campo na antiga Grécia. Suas fábulas são contadas pelos gregos que acompanharam suas peripécias pelo mundo matemático.

Euclides integrava o instituto de pesquisa estabelecido por Ptolomeu I em Alexandria e lá escreveu uma das obras mais importantes da história. *Os elementos*, obra de Euclides, estabelecia axiomas (verdades autoevidentes) da geometria, além de tratar das propriedades dos números, isto é, além dos volumes destinados à geometria, alguns dos volumes ocupavam-se de uma espécie de versão grega da Teoria dos Números. Assim, sobre a obra escrita pelo matemático, pode-se afirmar que:

> a parte central dos *Elementos* de Euclides lida com as propriedades dos números; nela se encontra o que talvez seja o primeiro momento brilhante do raciocínio matemático. Na proposição 20, Euclides explica uma verdade simples, porém fundamental, sobre os números primos: há um número infinito deles. A ideia começa pelo fato de que todo número pode ser gerado pela multiplicação de primos. (SAUTOY, 2007, p. 45)

Desta forma, ressaltamos que a primeira demonstração que garante a existência de infinitos primos deve-se a Euclides e foi formulada há cerca de 300 a.C. Euclides toma por base o argumento grego que diz que os primos são os blocos de construção de todos os outros números e questiona-se sobre a quantidade de blocos existentes.

Euclides acreditava não ser possível construir todos os números multiplicando as diferentes combinações de um bloco fixo de primos, pois pensou nos outros números que não poderiam ser gerados pelo bloco em questão. Não se sabe se a ideia foi do próprio Euclides ou se ele registou a ideia de algum pensador de Alexandria, mas o matemático demonstrou a maneira de construir primos que não fizessem parte de nenhuma lista fixa. Seu toque de mestre permitiu que ele pegasse uma lista de, por exemplo, 4 primos e adicionasse 1 ao seu produto. O número gerado pelo produto sempre deixava resto 1 quando dividido por qualquer um dos primos da lista, ou seja, o novo número era ele próprio primo ou era gerado por primos que não estivessem na lista fixada inicialmente. O Teorema de Euclides, bem como sua demonstração, é apresentado a seguir.

Teorema 2.1.1 (Teorema de Euclides) A quantidade de números primos é infinita.

Demonstração. Suponha que existe uma quantidade finita de números primos e denotemos estes por: $p_1, p_2, p_3, ..., p_k$. Consideremos o número $P = p_1 p_2 p_3 ... p_k + 1$ e chamemos de *p* o seu menor divisor primo. Obviamente, *p* não coincide com nenhum dos primos p_i, $1 \leq i \leq k$, pois, caso contrário, como ele divide *P*, teria que dividir a diferença $P - p_1 p_2 p_3 ... p_k = 1$, o que é impossível. Assim, *p* é um número primo que não pertence à sucessão $p_1, p_2, p_3, ..., p_k$ e, por consequência, a sucessão dada não pode formar o conjunto de todos os números primos. Logo, temos uma contradição à hipótese de que a quantidade de primos é finita.

Matemáticos como Goldbach e Euler também elaboraram suas demonstrações sobre a infinidade dos primos e, embora a demonstração de Euclides que acabamos de dar seja simples, ela não pode nos dar qualquer outra informação sobre o primo *p* apresentado, a não ser que ele seja, no máximo, menor ou igual ao número $P = p_1 p_2 p_3 ... p_k + 1$.

Além disso, é sabido que embora a sequência de primos seja, a princípio, razoavelmente suave, ela tem buracos de tamanhos distintos e irregulares. Isso nos leva a questionar sobre a existência de um número primo com um número dado qualquer de algarismos, sobre o tamanho dos buracos na representação dos primos etc.

Sobre o segundo questionamento, podemos notar na representação que existem primos que estão a uma distância pequena um do outro, por exemplo: (3,5), (5,7) e (11,13). Pares de números primos com essa propriedade são chamados de *primos gêmeos*. Uma das questões não resolvidas da matemática gira em torno desses curiosos pares de números: existem infinitos pares de números *primos gêmeos*?

Por outro lado, em contraste com esses pares de números primos muito próximos, podemos provar a existência de primos consecutivos arbitrariamente afastados, ou seja, podemos provar que os buracos entre os primos encontrados ficam cada vez maiores quando consideramos números maiores. Em algum lugar, existe uma sequência de 1000 números compostos consecutivos e bem mais adiante uma sequência bem maior de outros tantos números compostos consecutivos. Mostraremos, a seguir, que tal consideração é verdadeira.

Teorema 2.1.2 Para todo inteiro positivo *k*, existem *k* compostos consecutivos.

Demonstração. Seja $n = k + 1$ e considere os números $n!+2, n!+3, ..., n!+n$. A questão que devemos considerar é: algum desses números pode ser primo? A resposta é não. O primeiro número é par, pois $n!$ e 2 são pares. O segundo número é divisível por 3, pois $n!$ e 3 são

ambos divisíveis por 3 (assumindo que $n > 2$). De maneira análoga, concluímos que, em geral, $n!+i$ é divisível por *i*, para todo $i = 2, 3, ..., n$. Daí esses números não podem ser primos e, portanto, encontramos $n - 1 = k$ números compostos consecutivos.

Ainda sobre a temática da infinidade dos primos, Sautoy (2007, p. 47) destaca a forma maravilhosa que Euclides utilizou para argumentar, pois, apesar de não fazer ideia de como gerar primos explicitamente, o matemático conseguiu provar que eles nunca se esgotariam. Além disso, é incontestável que:

> os matemáticos tentaram, com diferentes graus de êxito, encontrar fórmulas que, mesmo sem gerar todos os números primos, produzissem ao menos uma lista de primos. Fermat acreditava haver encontrado uma. Ele supôs que elevando-se 2 à potência 2^N e adicionando-se 1, o número resultante $2^{2^N} + 1$ seria primo. Esse número é chamado de *n-ésimo número de Fermat*. (SAUTOY, 2007, p. 48)

Pierre de Fermat (1601-1665), jurista francês e estudioso que se dedicou à matemática como amador, é responsável por resultados e problemas que motivaram o extraordinário avanço da matemática. Após Euclides e Eratóstenes, é considerado um dos primeiros matemáticos a contribuir com os estudos teóricos da Teoria dos Números.

Figura 3

Fonte: Disponível em: < http://mathground.net/pierre-de-fermat-1601-1665/>. Acesso em: 25/02/2015.

Fermat (Figura 3) acreditava que os números denotados por ele, por meio da fórmula $F_n = 2^{2^n} + 1$, eram primos. Baseado em suas observações de que $F_1 = 5$, $F_2 = 17$ e $F_3 = 257$ e $F_4 = 65537$ são todos primos, Fermat escreveu, em 1640, uma carta endereçada a Frenicle, outro matemático amador, garantindo que os números com o formato descrito por

ele eram todos primos. Apesar de conseguir calcular o valor do quinto número, Fermat não tentou aplicar nenhum método de fatoração a F_5, confiando que o mesmo também era primo.

Mesmo trabalhando com fatoração, Frenicle não observou o erro de seu correspondente e parece ter concordado com a conjectura proposta. Desse modo, como nem Fermat e nem Frenicle observaram o erro da afirmação feita por Fermat em sua carta, coube a Euler, cem anos depois, a tarefa de mostrar a não primalidade de F_5. Como F_5 tem 10 algarismos, acredita-se que testar sua primalidade exigia ter à disposição ferramentas de que Fermat não dispunha.

Em 1732, Euler mostrou que o número F_5 de Fermat podia ser escrito da seguinte forma: $F_5 = 2^{2^5} + 1 = 4.294.967.297 = 641 \cdot 6700417$ e, portanto, não era primo e sim composto. Euler mostrou, ainda, que todo fator primo de números de Fermat (com $n \geq 2$) é da forma $k \cdot 2^{n+2} + 1$.

De acordo com Ribenboim (2012, p. 71), $F_4 = 65537$ é o maior número primo de Fermat conhecido e $F_{2478782}$, com o fator $3 \cdot 2^{2478785} + 1$ é o maior número de Fermat composto conhecido. Além disso, o autor destaca que "em fins de agosto de 2010, já eram conhecidos 243 números de FERMAT compostos" (RIBENBOIM, 2012, p. 71).

Como os números de Fermat crescem muito rapidamente, é uma tarefa complicada e laboriosa reconhecê-los como primos ou compostos. Ademais, a fatoração de números de Fermat compostos merece destaque, pois tem sido objeto de intensa investigação.

Além disso, há vários problemas em aberto sobre os primos de Fermat que despertam o interesse da comunidade matemática. Perguntas tais como: "Todo número F_n de Fermat com $n > 4$ é composto?", "Existem infinitos primos de Fermat?", e "Existem infinitos números de Fermat compostos?", permanecem sem respostas até hoje.

Os números de Fermat possuem uma longa história. Dado seu tamanho e a dificuldade de fatoração, esses números são ótimos para testar novos algoritmos de fatoração. Outra importância desses números reside no fato de que Gauss, em 1801, mostrou que, para um polígono regular de *n* lados ser construído com *régua e compasso*, é preciso que *n* seja um produto de uma potência de 2 por números primos de Fermat. De acordo com Sautoy (2007, p. 48):

> Gauss tinha uma afeição especial pelos números de Fermat. O fato de que 17 seja um dos primos de Fermat é fundamental para entender por que Gauss conseguiu

> construir seu polígono perfeito de 17 lados. No grande tratado *Disquisitiones Arithmeticae*, Gauss demonstra por que, se o *N*-ésimo número de Fermat for primo, é possível fazer a construção geométrica de um polígono de *N* lados usando apenas uma régua e um compasso.

Apesar de Fermat só ter conseguido gerar quatro primos, o matemático amador foi bastante competente na descoberta de propriedades especiais sobre tais números. Fermat descobriu, por exemplo, que os números primos que deixam resto 1 quando divididos por 4 podem ser expressos pela soma de dois quadrados. Em 1640, Fermat registrou a descoberta em uma carta que enviou a seu contemporâneo Marin Mersenne (Figura 4) com quem se correspondia em uma troca de conhecimentos matemáticos.

Marin Mersenne (1588-1648) foi matemático, filósofo e teórico musical. Conhecido pelo estudo dos primos de Mersenne, teve papel de destaque na difusão da ciência devido à sua extensa correspondência com cientistas da época. Nascido na cidade de Maine e grande influenciador das ciências, em especial da matemática francesa nos séculos XVI e XVII, o monge Mersenne teve entre seus correspondentes, além de Fermat, nomes ilustres como Descartes, Pascal e Galileu.

Figura 4

Fonte: Disponível em: < http://extern.peoplecheck.de/link.php?q=marin+mersenne&url=http%3A%2F%2Fgogh-creative.deviantart.com%2Fart%2FMarin-Mersenne-375455893>. Acesso em 25/02/2015.

Os números que recebem o nome de *números de Mersenne* são da forma $M_p = 2^p - 1$, onde p é um número primo. São primos de Mersenne os números $M_2 = 3$, $M_3 = 7$, $M_5 = 31$, $M_7 = 127$ etc. No intervalo $2 \le p \le 1000$, os números de Mersenne que são primos, chamados de *primos de Mersenne*, correspondem aos seguintes valores de *p*: 2, 3, 5, 7, 13, 19, 31, 61, 89, 107, 127, 521 e 607.

Determinar quais números de Mersenne são primos é mais uma questão legada dos gregos. Ribenboim (2012, p. 73), relata que:

> desde o tempo de MERSENNE, era sabido que certos números de MERSENNE são primos e que outros são compostos. Por exemplo $M_2 = 3$, $M_3 = 7$, $M_5 = 31$, $M_7 = 127$ são primos, enquanto que $M_{11} = 23 \cdot 89$. Em 1640, MERSENNE afirmou que M_q é primo para q = 13, 17, 19, 31, 67, 127 e 257; estava ele enganado em relação a 67 e 257; também não incluíra 61, 89 e 107 (entre os números inferiores a 257) que também fornecem números de MERSENNE primos. Sua afirmação era extraordinária, em face da grandeza dos números envolvidos.

É possível que os conhecimentos musicais de Mersenne o tenham ajudado na descoberta de sua fórmula. Até hoje (junho de 2015) são conhecidos 48 números de Mersenne que são primos e o maior deles foi descoberto em 25 de janeiro de 2013, possuindo em seu sistema decimal quase 13 milhões de algarismos.

A busca pelos primos, um dos mais fascinantes temas matemáticos, esconde velhos mistérios e é cercada de inúmeras e desafiadoras possibilidades, tanto que no século XVIII Euler surge como um nome importante no cenário matemático da época.

Figura 5

Fonte: Disponível em: < http://pt.wikipedia.org/wiki/Leonhard_Euler>. Acesso em: 25/02/2015.

Leonhard Euler (Figura 5), nascido em 1707, foi um grande matemático e físico suíço. O pai, clérigo, esperava que Euler se unisse à Igreja, mas isso não foi possível, pois o matemático logo chamou a atenção de poderosos e, de acordo com Sautoy (2007, p. 50), viu-se adulado pelas academias de toda a Europa.

Euler se interessava por diversas áreas: hidráulica, balística, construção de navios, matemática e passou até pela música. Sua produção era de uma extensão tão vasta que mesmo

tempos depois de sua morte, muitos trabalhos inéditos foram encontrados em seus arquivos. Além disso:

> a paixão de Euler pela teoria dos números foi estimulada pela correspondência com Christian Goldbach, um matemático amador alemão que viveu em Moscou e estava empregado extraoficialmente como secretário da Academia de Ciências de São Petersburgo. Como o matemático amador Mersenne, Goldbach era fascinado por brincadeiras e experimentos numéricos. Foi a Euler que Goldbach comunicou sua conjectura de que todo número par poderia ser expresso como a soma de dois primos. Em troca, Euler pedia a Goldbach que testasse as diversas provas que desenvolvia para confirmar o misterioso catálogo de descobertas de Fermat. (SAUTOY, 2007, p. 53-54)

A relação de Euler com os números primos era muito boa, mesmo porque ele já tentava há algum tempo provar as afirmações de Fermat, que, por sua vez, não tinha tanta preocupação em provar. Seja por alegar o tamanho da prova para o papel que dispunha ou a simplicidade da prova em questão, Fermat não estava tão preocupado em registrar demonstrações. Euler, ao contrário, tinha paixão pela prova apesar de ser um matemático experimental. Euler era um exímio elaborador de tabelas de primos e produziu tabelas com todos os primos até mais de 100.000.

De acordo com Sautoy (2007, p. 54), uma de suas descobertas mais curiosas foi uma fórmula que parecia gerar um número inusitado de primos. Euler inseriu na fórmula $x^2 + x + 41$ os números de 0 a 39, obtendo uma lista de primos. É claro que Euler acreditava que em algum momento a fórmula se mostraria falha. Para o número 41, por exemplo, isso é óbvio. Mesmo assim:

> [...] Euler ficou bastante impressionado com a capacidade da fórmula de gerar tantos primos. Ele ponderou quais outros números serviriam, além de 41, e descobriu que para q = 2, 3, 5, 11 e 17 a fórmula $x^2 + x + q$ também emitia primos quando alimentada com números de 0 a $q - 2$. (SAUTOY, 2007, p. 54)

Euler, que tanto trabalhou em tabelas de primos, chegou à conclusão que existem mistérios que a mente humana jamais penetrará e a organização dos primos é um deles. Apesar desse pensamento, Euler foi incansável na busca de uma fórmula que decifrasse, pelo menos, parte do mistério dos primos. Assim, tempos depois, definiu uma função denotada pela letra grega ζ (zeta). A função, definida por $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} + \ldots$, foi de grande importância para as descobertas futuras. Sautoy (2007, p. 55) alerta para o fato de que:

> posteriormente, ficaria claro que Euler tinha em mãos uma equação que romperia o impasse dos primos, mas seriam necessários outros cem anos, e outra grande mente, para demonstrar o que Euler não percebera. Essa mente pertencia a Bernhard Riemann. Entretanto, Gauss foi o responsável em inspirar a nova perspectiva de Riemann, introduzindo outro de seus clássicos passos laterais.

Carl Friedrich Gauss (Figura 6) trouxe valorosas contribuições e tentou atacar o problema dos primos de um ponto de vista diferente. Nascido em 1777, foi matemático, astrônomo e físico alemão. Sua contribuição e influência em diversas áreas matemáticas o fez ficar conhecido entre alguns matemáticos como *o príncipe da matemática*. Aos 16 anos, teve o vislumbre de uma geometria diferente da de Euclides e um ano depois começou uma busca crítica pelas provas na Teoria dos Números, além de tentar preencher as lacunas que outros matemáticos deixaram.

Figura 6

Fonte: Disponível em: < http://carlgaussmatematico.blogspot.com.br/2011/11/biografia-carl-gauss.html>. Acesso em: 25/02/2015.

Gauss acreditava que se não era possível determinar uma fórmula para os primos, valia a pena tentar encontrar uma fórmula que contabilizasse a quantidade de primos em um intervalo de 1 a *n*, por exemplo. O olhar diferenciado de Gauss sobre os números primos foi influenciado por um presente: um livro de logaritmos. Com 15 anos de idade, um ano depois de ter ganhado o livro, Gauss já pensava em uma nova abordagem para tratar dos primos. Sautoy (2007, p. 55) conta que na contracapa do livro de logaritmos de Gauss havia uma tabela de primos. Para Gauss, a presença dos dois conteúdos em um mesmo livro era, no mínimo, curiosa. Apesar das tabelas de primos serem consideradas inúteis,

> as tabelas de logaritmos ajudaram a acelerar o mundo do comércio e da navegação, que florescia no século XVII. Graças ao diálogo que criavam entre a multiplicação e a adição, as tabelas de logaritmos facilitaram a resolução de problemas complicados, que envolviam a multiplicação de dois grandes números, transformando-os na simples adição de seus logaritmos. (SAUTOY, 2007, p. 56)

As tabelas de logaritmos, como bem destacou o autor, eram muito importantes na época. As tabelas de primos, porém, não passavam de curiosidade. Enquanto que os logaritmos seguiam o padrão esperado, os números primos eram totalmente aleatórios. Não se sabia o que prever sobre os primos, nem mesmo qual era o próximo primo depois de um dado número. Como não havia uma fórmula que determinasse primos, sua descoberta sempre foi deveras trabalhosa.

A esse respeito Gauss resolveu fazer uma pergunta diferente. Ao invés de preocupar-se com uma fórmula que pudesse ordenar os primos, o matemático, como já mencionado, resolveu procurar a quantidade de primos entre os primeiros 100 ou os primeiros 1000 números, por exemplo. Assim:

> se tomássemos o número N, haveria alguma maneira de estimar quantos primos encontraríamos entre os números 1 e N? Por exemplo, existem 25 primos até o número 100. Portanto, temos uma chance de um em quatro de encontrar um primo se escolhermos um número aleatório entre 1 e 100. Como se altera essa proporção se buscarmos os primos de 1 a 1.000 ou de 1 a 1.000.000? (SAUTOY, 2007, p. 56-57)

A tabela a seguir, que indica a quantidade de primos de 1 até determinada potência de 10, apresenta a regularidade observada pelo matemático.

Tabela 1 – Quantidade de primos em potências de 10

N	Número de primos de 1 a N, frequentemente chamado de $\pi(N)$	Em média, quantos números precisamos contar até atingir um número primo
10	4	2,5
100	25	4,0
1.000	168	6,0
10.000	1.229	8,1
100.000	9.592	10,4
1.000.000	78.498	12,7
10.000.000	664.579	15,0
100.000.000	5.761.455	17,4
1.000.000.000	50.847.534	19,7
10.000.000.000	455.052.511	22,0

Fonte: SAUTOY, 2007, p. 57

Consequentemente, uma vez que estava:

> armado com tabelas de números primos, Gauss iniciou sua busca. Ao observar a proporção de primos no universo de números, notou o surgimento de um padrão à medida que a contagem se elevava. Apesar da aleatoriedade desses números, parecia ser possível entrever uma regularidade estonteante. (SAUTOY, 2007, p. 57)

Gauss percebeu uma relação entre a multiplicação e a adição nos primos e essa relação era justamente aquela representada pelos logaritmos. Segundo os seus cálculos, a cada multiplicação por 10, a proporção de primos aumentava 2,3 e não 1 como acontecia nos logaritmos de base decimal, ou seja, os primos seguiam logaritmos que não eram de base 10. Desse modo:

> a descoberta de Gauss foi o fato de que os primos podem ser contados usando-se logaritmos cuja base é um número especial, chamado *e*, que, com 12 casas decimais, tem o valor de 2.718 281 828 459... (da mesma forma que π, esse número possui uma expansão decimal infinita sem padrões repetitivos) [...] A tabela que Gauss criou aos 15 anos de idade o levou à seguinte conjectura: entre os números 1 a *N*, aproximadamente 1 em cada log(*N*) será primo (onde log(*N*) denota o logaritmo de *N* na base *e*). (SAUTOY 2007, p. 58)

Gauss estimou que o número de primos entre 1 e *N* era, aproximadamente, $\frac{N}{\log N}$. Gauss não acreditava que a fórmula fornecia o número exato de primos, mas ponderou que a estimativa era, no mínimo, razoável.

Embalado por um pensamento que nenhum outro matemático jamais tivera, Gauss deu um passo atrás, mas fez uma descoberta importantíssima no que concerne aos números primos. A partir dos estudos de Gauss, "surgiu o costume de expressar o número de primos que encontramos entre os números 1 a *N* através do símbolo $\pi(N)$ [...]" (SAUTOY, 2007, p. 59).

Gauss estava diante de uma descoberta incrível, mas manteve-se relutante, pois apesar dos indícios sobre a conexão de números primos e logaritmos, nada garantia que esse padrão iria permanecer em contagens mais altas. Gauss não era dado a especulações como muitos matemáticos, queria anunciar sua descoberta desde que tivesse em mãos a prova definitiva de que ela estava correta, que não ia falhar em um ponto do caminho.

A preocupação de Gauss era tanta que ele chegou a criptografar alguns de seus resultados usando uma linguagem própria. Se ele não podia provar, tampouco poderia divulgar descobertas incompletas. Apesar disso, a descoberta de Gauss sobre os primos

mostrou-se verdadeira, mesmo depois de contestada por Legendre, que acreditava faltar um acréscimo na fórmula apresentada por Gauss.

Tempos depois da descoberta de Gauss, mais ou menos em 1810, a matemática passa a fazer parte do currículo dos novos ginásios e universidades. Os alunos eram estimulados a estudar matemática não apenas como apoio às ciências e sim como uma disciplina própria. A revolução educacional, iniciada na Alemanha, teve um grande impacto na compreensão dos matemáticos sobre sua própria disciplina e, de acordo com Sautoy (2007, p. 69): "o estudo dos números primos foi particularmente revolucionado".

Um dos expoentes dessa revolução foi Augustin-Louis Cauchy (Figura 7). Matemático francês, Cauchy demonstrou desde cedo talento para matemática. Um amigo de seu pai, o matemático Lagrange, reconheceu o talento prodigioso do menino e aconselhou seu pai a deixá-lo longe do mundo matemático até seus 17 anos. Um conselho valoroso, já que ao voltar, foi responsável por grandes feitos e, em especial, pelo primeiro avanço da matemática moderna: a introdução do rigor na análise matemática.

De acordo com Sautoy (2007, p. 75): "Cauchy teve problemas com as autoridades de Paris, por desviar os estudantes das aplicações práticas da matemática", sendo considerado um dos únicos que fazia uso da chamada "matemática pura".

Cauchy também empenhou seus esforços no estudo dos números imaginários. Ele e outros matemáticos queriam saber o que iria acontecer se estendessem o conceito de função a esse novo conjunto de números.

Figura 7

Fonte: Disponível em: < http://fan-people.com/augustin-cauchy-photo2/>. Acesso em: 25/02/2015.

Bernhard Riemann corroborava com as ideias de Cauchy, tanto que descreveu sua dissertação baseada no estudo dos números imaginários. Até então, Riemann não havia dado atenção ao estudo dos números primos.

Georg Friedrich Bernhard Riemann (Figura 8), nascido em 1826, foi um matemático alemão que apresentou contribuições fundamentais à análise matemática e à geometria diferencial. Quando criança, precisou lutar contra um perfeccionismo incapacitante. Quase não entregava os trabalhos de escola, pois esses deveriam estar não menos que perfeitos. Tímido, acabou refugiando-se na matemática onde fez descobertas grandiosas.

Figura 8

Fonte: Disponível em: < http://sites.middlebury.edu/fyse1229hunsicker/biography/>. Acesso em: 25/02/2015.

Após a morte de Gauss, um de seus grandes admiradores, o matemático alemão Johann Peter Gustav Lejeune Dirichlet, assumiu sua cadeira na Universidade de Göttingen e Riemann acabou ganhando uma espécie de mentor intelectual com quem podia, casualmente, conversar e trocar figurinhas matemáticas.

Riemann, influenciado pelas ideias de Dirichlet, passou a demonstrar interesse pela função zeta. Dirichlet usou essa função para demonstrar uma das ideias de Fermat e Riemann, que ainda estava ocupado com o estudo dos números imaginários, só conseguia pensar que essa era uma função interessante na qual poderia trabalhar com tais números. Algum tempo depois, tendo percebido melhor os desdobramentos da função zeta, Riemann se viu inserido no terreno dos números primos. O matemático constatou que a função zeta poderia ajudá-lo a provar que a estimativa de Gauss para os números primos estava, de fato, correta. É importante lembrar que:

> as descobertas de Riemann foram muito além dessa ideia isolada. Ele se viu observando os primos a partir de uma perspectiva completamente diferente. A função zeta passou a tocar uma música que tinha potencial de revelar os segredos dos primos. [...] Em novembro de 1859, Riemann expôs suas descobertas em um artigo publicado no periódico mensal da Academia de Berlim. Essas dez páginas de densa matemática foram as únicas que Riemann publicou, em toda sua vida, sobre os números primos, mas o artigo teria um efeito fundamental sobre a maneira como eram percebidos. (SAUTOY, 2007, p. 92)

O fato de Riemann ter escrito um artigo cheio de brechas foi, no mínimo, intrigante. Riemann mencionava suas limitações a respeito de sua hipótese e afirmava que muitas das declarações feitas vinham de resultados que ele acreditava não estar prontos para ser divulgados. Em seus estudos:

> a conexão que Riemann encontrou os números primos e os pontos no nível do mar na paisagem zeta não poderia ser mais direta. Gauss tentara estimar quantos primos havia do número 1 a qualquer número *N*. Riemann, entretanto, conseguiu produzir uma fórmula *exata* para o número de primos até *N*, usando as coordenadas desses zeros. (SAUTOY, 2007, p. 99)

A nova função estudada por Riemann aperfeiçoava a função de Gauss e apresentava um nível maior de precisão, ou seja, os desdobramentos da função zeta para números complexos, estudado por Riemann, o levaram a acreditar que a estimativa de Gauss para a quantidade de primos em um dado intervalo era, de fato, correta.

Com os notáveis progressos feitos pelo matemático ao formular a chamada *Hipótese de Riemann* e relacioná-la com o problema da contagem da quantidade de primos proposto por Gauss, Sautoy (2007, p. 101) acredita que "Riemann descobriu o cálice sagrado que Gauss havia buscado: a fórmula exata para contar o número de primos até *N*". Entretanto, apesar da afirmação feita pelo autor, estudos comprovam que o resultado acabou por ser provado (usando métodos poderosos em análise complexa), de maneira independente, por Jacques Hadamard e Charles De la Vallée Poussin em 1896.

O estudo dos primos mostra-se até hoje desafiador e deveras intrigante. Além da hipótese de Riemann, muitas questões, mesmo nos dias de hoje, apresentam-se sem respostas. Sautoy (2007, p.18) destaca: "os números primos ocupam lugar tão fundamental na matemática que qualquer progresso na compreensão de sua natureza terá um impacto grandioso". O próximo tópico encarrega-se de apresentar o resultado que trouxe um desses impactos grandiosos na época em que foi apresentado: o *Teorema Fundamental da Aritmética*.

2.3 O Teorema Fundamental da Aritmética

Vimos que um número primo possui apenas dois divisores, o número 1 e o próprio primo, portanto, não pode ser escrito como um produto de fatores menores do que ele. Por exemplo, como 13 é primo, a única forma de escrevê-lo, a menos da ordem dos fatores,

como um produto é $13 = 1 \cdot 13$. Já o número 30, composto, possui divisores não triviais, podendo ser *fatorado* como produto de números menores, a saber: $30 = 2 \cdot 15 = 3 \cdot 10 = 5 \cdot 6$. Observe que nenhuma das fatorações apresentadas envolve apenas primos. A única fatoração que envolve apenas primos e que não foi escrita anteriormente é $30 = 2 \cdot 3 \cdot 5$.

Os números primos têm a importante propriedade de formar todos os outros números, ou seja, todos os números podem ser escritos como um produto de primos como fizemos acima com o exemplo particular. De acordo com Sautoy (2007, p. 13):

> esses números são os próprios átomos da aritmética. São os números indivisíveis, que não podem ser representados pela multiplicação de dois números menores. Os números 13 e 17 são primos, ao contrário de 15, que pode ser expresso como 3 vezes 5. Os primos são as pérolas que adornam a vastidão infinita do universo de números que os matemáticos exploraram ao longo dos séculos. Eles despertam a admiração dos matemáticos: 2, 3, 5, 7, 11, 13, 17, 19, 23, ... – números eternos que existem em uma espécie de mundo independente de nossa realidade física. São um presente da natureza para os matemáticos.

O conjunto dos números naturais é fechado para a adição e para a multiplicação e, do ponto de vista multiplicativo, os números primos são os mais simples e, como enunciamos, suficientes para gerar todos os outros números. Tal enunciado é conhecido como *Teorema Fundamental da* Aritmética e é o objeto de estudo do tópico em questão. Porém, antes de enunciar o teorema e de demonstrá-lo, vamos conhecer um resultado que poderá tornar sua demonstração mais clara.

Lema 2.4.1 Se *n* é um número composto, então o menor divisor próprio de *n* é primo.

Demonstração. Seja *d* o menor divisor próprio de *n*, com $d \neq 1$. Se *d* fosse composto, possuiria um divisor próprio, digamos d_1. Mas, $d_1 \mid d$ e $d \mid n$, implicando que $d_1 \mid n$ e, como, $d_1 < d$, haverá um divisor próprio de *n* menor que *d*, contrariando a sua minimalidade. Logo, o menor divisor próprio de *n* deve ser primo.

Finalmente, chegamos ao ponto principal do tópico. Vejamos, a seguir, o enunciado e a demonstração do teorema.

Teorema 2.4.1 (Teorema Fundamental da Aritmética) Todo número natural maior do que 1 ou é primo ou se escreve de modo único (a menos da ordem de fatores) como um produto de números primos.

Demonstração. Se *n* é primo, nada temos a demonstrar. Suponhamos, então, que *n* seja composto. Pelo lema anterior, o menor divisor próprio de *n* é primo. Chamando-o de p_1, podemos escrever $n = p_1 n_1$. Se n_1 é primo, o resultado vale imediatamente. Se, ao contrário, n_1 é composto, seu menor divisor próprio é primo e, chamando-o de p_2, podemos escrever $n = p_1 n_1 = p_1 p_2 n_2$. O processo pode ser repetido e como a cada passo $n_i < n_{i-1}$, ou seja, forma-se uma sequência decrescente de naturais maiores que 1 e haverá um momento no qual teremos $n_m = p_m$ primo e, assim, $n = p_1 \cdot \ldots \cdot p_m$. Como os primos obtidos não são necessariamente distintos, podemos contar a quantidade que cada primo p_i aparece. Chamando essa quantidade de α_i, teremos: $n = p_1^{\alpha_1} \cdot \ldots \cdot p_m^{\alpha_m}$.

Para provar a unicidade da escrita, vamos supor que haja duas maneiras de escrever o natural *n*, ou seja, $n = p_1^{\alpha_1} \cdot \ldots \cdot p_m^{\alpha_m} = q_1^{\beta_1} \cdot \ldots \cdot q_r^{\beta_r}$. Como p_1 é primo e divide $n = q_1^{\beta_1} \cdot \ldots \cdot q_r^{\beta_r}$, pela definição de primos, $p_1 \mid q_i$ para algum *i* e, como eles são primos, devem ser iguais e, assim, podemos reordenar os primos da segunda decomposição para que $p_1 = q_1$. Usando argumento semelhante, podemos concluir que os expoentes devem ser iguais também. Da igualdade $p_1^{\alpha_1} \cdot \ldots \cdot p_m^{\alpha_m} = q_1^{\beta_1} \cdot \ldots \cdot q_r^{\beta_r}$ e, uma vez que, $p_1 = q_1$ e $\alpha_1 = \beta_1$, podemos concluir que $p_2^{\alpha_2} \cdot \ldots \cdot p_m^{\alpha_m} = q_2^{\beta_2} \cdot \ldots \cdot q_r^{\beta_r}$. De modo análogo, quando se esgotarem os primos p_i também se esgotarão os primos q_i. Assim, a quantidade de primos nas duas decomposições é a mesma e, a menos da ordem dos fatores, elas possuem os mesmos fatores com os mesmos expoentes.

Uma vez inseridos no histórico dos primos, vamos destacar no próximo capítulo outros tópicos bastante importantes da Teoria dos Números, que nos ajudarão a compreender, posteriormente, o universo da criptografia.

3 CAPÍTULO II – TÓPICOS DE TEORIA DOS NÚMEROS

A teoria dos números é o ramo da matemática que se ocupa do estudo dos números inteiros, procurando explicar sua origem, as relações entre eles e as suas propriedades. Embora, como o próprio nome indica, trate-se de teoria, é possível relacioná-la a outras ciências e conteúdos práticos como a criptografia.

Os sistemas de codificação e decodificação, essenciais ao estudo da criptografia, são desenvolvidos através de bases matemáticas seguras que permitem a confidencialidade da informação. Nesse sentido, vamos explorar no capítulo alguns tópicos relevantes da teoria dos números, que serão de grande ajuda na compreensão das técnicas matemáticas empregadas no desenvolvimento da criptografia.

3.1 Divisibilidade

Definição 3.1.1 Dados os números $a,b \in Z$, com $a \neq 0$, dizemos que *a* divide *b*, e escrevemos $a \mid b$, se existir um inteiro *n* tal que $b = an$, ou seja, $a \mid b \Leftrightarrow \exists n \in Z; b = an$. Caso *a* não divida *b*, escrevemos $a \nmid b$.

Quando $a \mid b$, dizemos também que *a* é um divisor de *b* ou, equivalentemente, que *b* é um múltiplo de *a*. Nesse sentido, podemos observar, por exemplo, que 5 é um divisor de 30 ou que 30 é um múltiplo de 5, pois $5 \mid 30$, ou seja, $30 = 5 \cdot 6$. De modo análogo, podemos concluir que 12 não é um múltiplo de 7, pois não existe nenhum inteiro *n* tal que $12 = 7n$, ou seja, temos que $7 \nmid 12$. A seguir, apresentaremos algumas propriedades fundamentais para a divisibilidade.

(1) $a \mid a$, para todo $a \in Z$.

(2) Se $a \mid b$ e $b \mid c$, então $a \mid c$, para todo $a,b \in Z$.

(3) Se $a \mid b$ e $c \mid d$, então $ac \mid bd$.

(4) Se $a \mid b$, então $a \mid mb$, para todo $m \in Z$.

(5) Se $a \mid b$ e $a \mid c$, então $a \mid (bx + cy), \forall x, y \in Z$.

(6) Se $a \mid b$ e $b \mid a$, então $a = \pm b$.

3.2 Divisão de inteiros

Teorema 3.2.1 (Algoritmo da Divisão) Dados os números inteiros positivos *b* e *a*, existe um único par de números inteiros *q* e *r*, denominados, respectivamente, de *quociente* e *resto* da divisão de *b* por *a*, tais que: $b = aq + r$ e $0 \leq r < a$.

Assim, dividir *b* (chamado dividendo) por *a* (o divisor) significa encontrar o quociente e o resto que satisfazem ao teorema anterior. Dado que $17 = 5 \cdot 3 + 2$ e $2 < 5$, podemos dizer que a divisão de 17 por 5 apresenta quociente 3 e resto 2. Porém, apesar de $23 = 7 \cdot 2 + 9$, não podemos dizer que a divisão de 23 por 7 deixa quociente 2 e resto 9, pois devemos lembrar que o resto não pode ser maior que o divisor.

Além disso, sempre que $a \mid b$, o resto da divisão de *b* por *a* é zero, ou seja, a divisão é *exata*. É o caso, por exemplo, da divisão de 20 por 4, onde obtemos quociente 5 e resto 0, ou seja, $20 = 4 \cdot 5 + 0$.

3.3 Máximo Divisor Comum – MDC

Definição 3.3.1 Sejam *a* e *b* dois inteiros não simultaneamente nulos, isto é, $a \neq 0$ ou $b \neq 0$. Chama-se *máximo divisor comum* de *a* e *b*, o inteiro positivo *d* que satisfaz as condições:

1. $d \mid a$ e $d \mid b$, ou seja, *d* é um divisor comum de *a* e de *b*;
2. Se *c* é um inteiro tal que $c \mid a$ e $c \mid b$, então $c \leq d$, ou seja, *d* é o maior dos divisores comuns de *a* e de *b*.

Utilizaremos a notação $d = mdc(a,b)$ para indicar que *d* é o máximo divisor comum entre *a* e *b*. Além disso, vamos considerar $D(a)$ como o conjunto dos divisores positivos de *a*. A partir disso teremos, por exemplo, que como $D(20) = \{1,2,4,5,10,20\}$ e $D(15) = \{1,3,5,15\}$ é possível concluir que $mdc(20,15) = 5$.

Além disso, vale ressaltar que como o número 1 é divisor de qualquer inteiro, o máximo divisor comum de dois inteiros *a* e *b* nunca é vazio, sendo limitado por *a* e por *b*. Ademais, como podemos garantir a existência de um elemento máximo, podemos concluir que o máximo divisor comum sempre existe. A seguir, apresentaremos um teorema e um corolário que apontam propriedades fundamentais para o estudo dos números primos.

Teorema 3.3.1 Se um número primo *p* não divide um inteiro *a*, então *a* e *p* são relativamente primos ou primos entre si.

Demonstração. Seja $d = mdc(a, p)$, então $d \mid a$ e $d \mid p$. Da relação $d \mid p$, temos que $d = 1$ ou $d = p$, pois *p* é primo. Como a segunda igualdade é impossível, pois *p* não divide *a*, então $d = 1$, ou seja, $mdc(a, p) = 1$. Logo, *a* e *p* são relativamente primos.

Corolário 3.3.1 (Propriedade Fundamental dos Números Primos) Se *p* é um primo tal que $p \mid ab$, então $p \mid a$ ou $p \mid b$.

Demonstração. Se $p \mid a$, não há o que demonstrar. Se, do contrário, $p \nmid a$, então pelo teorema anterior, $mdc(p, a) = 1$ e, assim, $p \mid b$.

A seguir, demonstraremos um método prático para determinação do MDC. O método, denominado Algoritmo de Euclides, também conhecido como *método das divisões sucessivas,* nos permite a utilização de uma série finita de divisões que nos possibilitam a determinação do MDC. Antes, porém, apresentaremos um Lema que poderá ajudar no entendimento do algoritmo.

Lema 3.3.1 Sejam *a* e *b* inteiros, com $b \neq 0$, e sejam *q* e *r* o quociente e o resto da divisão de *a* por *b*, respectivamente, ou seja, $a = bq + r$. Então, $mdc(a, b) = mdc(b, r)$.

Demonstração. Seja $d = mdc(a, b)$. Daí, $d \mid a$ e $d \mid b$, o que implica $d \mid a - bq$, isto é, $d \mid r$. Logo, $d \mid b$ e $d \mid r$. Dessa maneira, $d \leq mdc(b, r) = d'$ (I). Reciprocamente, $d' \mid b$ e $d' \mid r$, o que implica $d' \mid bq + r$. Daí, temos que $d' \mid a$ e $d' \mid b$. Portanto, $d' \leq d$ (II). De (I) e (II), concluímos que $d = d'$.

Para determinar o máximo divisor comum de 400 e 148, por exemplo, podemos dividir 400 por 148 e observar o resto da divisão. Como $400 = 2 \cdot 148 + 104$, temos, pelo Lema anterior, que $mdc(400,148) = mdc(148,104)$. Portanto, devemos proceder realizando divisões sucessivas entre o divisor e o resto até o ponto em que o resto em questão é zero, ou seja, até encontrar uma divisão exata. Quando isso ocorre, o divisor é o próprio MDC.

Nesse sentido, apresentaremos todos os passos para o exemplo acima, após a exposição e demonstração do teorema a seguir.

Teorema 3.3.1 (Algoritmo de Euclides) Sejam $r_0 = a$ e $r_1 = b$ inteiros não negativos com $b \neq 0$. Se o algoritmo da divisão for aplicado sucessivamente para se obter $r_j = q_{j+1} r_{j+1} + r_{j+2}$, $0 \leq r_{j+2} < r_{j+1}$ para $j = 0,1,2,...,n-1$ e $r_{n+1} = 0$, então $mdc(a,b) = r_n$, o último resto não nulo das divisões.

Demonstração. Vamos aplicar o teorema 3.2.1 para dividir $r_0 = a$ e $r_1 = b$, obtendo $a = q_1 b + r_2$, ou seja, $r_0 = q_1 r_1 + r_2$. Em seguida, dividimos r_1 e r_2, obtendo $r_1 = q_2 r_2 + r_3$ e assim, sucessivamente, até a obtenção do resto $r_{n+1} = 0$. Como, a cada passo, o resto é sempre menor que o anterior, e estamos lidando com inteiros positivos, é claro que após um número finito de aplicações do algoritmo da divisão, teremos resto nulo. Temos, pois, as seguintes sequências de equações:

$$\begin{aligned}
r_0 &= q_1 r_1 + r_2, && 0 < r_2 < r_1 \\
r_1 &= q_2 r_2 + r_3, && 0 < r_3 < r_2 \\
r_2 &= q_3 r_3 + r_4, && 0 < r_4 < r_3 \\
&\ldots \\
r_{n-2} &= q_{n-1} r_{n-1} + r_n, && 0 < r_n < r_{n-1} \\
r_{n-1} &= q_n r_n + 0
\end{aligned}$$

A última destas equações nos diz que o máximo divisor comum de r_n e r_{n-1} é r_n. A penúltima diz que esse número é igual a $mdc(r_{n-1}, r_{n-2})$ e, prosseguindo desta maneira teremos, por sucessivas repetições do Lema 3.3.1, a sequência:

$$r_n = mdc(r_{n-1}, r_n) = mdc(r_{n-2}, r_{n-1}) = \ldots = mdc(r_1, r_2) = mdc(r_0, r_1) = mdc(a,b)$$

Portanto, o máximo divisor comum de *a* e *b* é o último resto não nulo da sequência de divisões descrita.

□

O processo de determinação do máximo divisor comum pelo Algoritmo de Euclides, nos permite trabalhar com números cada vez menores, o que facilita as divisões.

Logo, para determinar o MDC de 400 e 148, exemplo apresentado após o Lema 3.3.1, procedemos da forma indicada no algoritmo. Assim:

$$400 = 2 \cdot 148 + 104 \qquad mdc(400,148) = mdc(148,104)$$
$$148 = 1 \cdot 104 + 44 \qquad mdc(148,104) = mdc(104,44)$$
$$104 = 2 \cdot 44 + 16 \qquad mdc(104,44) = mdc(44,16)$$
$$44 = 2 \cdot 16 + 12 \qquad mdc(44,16) = mdc(16,12)$$
$$16 = 1 \cdot 12 + 4 \qquad mdc(16,12) = mdc(12,4)$$
$$12 = 3 \cdot 4 + 0 \qquad mdc(12,4) = 4$$

O Algoritmo de Euclides também pode ser realizado utilizando o diagrama abaixo. É claro que a ideia é a mesma e o algoritmo não é modificado de modo algum. O diagrama talvez nos permita até enxergar de maneira mais simples.

	q_1	q_2	q_3	$\cdots$	q_{n-1}	q_n	q_{n+1}
b	a	r_1	r_2	$\cdots$	r_{n-2}	r_{n-1}	$r_n = (a, b)$
r_1	r_2	r_3	r_4	$\cdots$	r_n		

É importante observar que decorre também do Algoritmo de Euclides que se *d* é o máximo divisor comum entre *a* e *b*, *d* pode ser escrito como combinação linear de *a* e *b*, ou seja, o algoritmo nos fornece um meio prático de escrever o MDC de dois números como soma de dois múltiplos dos números em questão. Assim, por exemplo, a equação $6x + 10y = 1$ não tem raízes inteiras, pois 6 e 10 não são primos entre si. Por outro lado, a equação $372x + 162y = 6$ possui raízes inteiras, pois 6 é MDC entre 372 e 162.

	2	3	2	1	2
372	162	48	18	12	6
48	18	12	6		

3.4 Congruências

Definição 3.4.1 Diremos que dois números naturais *a* e *b* são *congruentes módulo n*, e representamos por $a \equiv b(\mathrm{mod}\, n)$ quando $a-b$ é um múltiplo de *n*, ou seja: $a \equiv b(\mathrm{mod}\, n) \Leftrightarrow n \mid a-b$.

Por exemplo, $29 \equiv 11(\mathrm{mod}\, 3)$, pois $29-11=18$ é um múltiplo de 3. Além disso, a congruência $a \equiv b(\mathrm{mod}\, n)$ também acontece se os restos da divisão euclidiana de *a* e *b* por *n* são iguais. Quando a relação $a \equiv b(\mathrm{mod}\, n)$ for falsa, dizemos que os números não são congruentes ou que são *incongruentes*. Escrevemos, nesse caso, $a \not\equiv b(\mathrm{mod}\, n)$. Os números 31 e 23 não são congruentes módulo 5, pois não deixam o mesmo resto na divisão por 5. Daí, $31 \not\equiv 23(\mathrm{mod}\, 5)$.

Assim sendo, dado um inteiro positivo *n*, a relação de congruência módulo *n* satisfaz as seguintes propriedades:

(1) $a \equiv a(\mathrm{mod}\, n)$ para qualquer inteiro *a*, ou seja, a relação é *reflexiva*.
(2) Se $a \equiv b(\mathrm{mod}\, n)$, então $b \equiv a(\mathrm{mod}\, n)$ para quaisquer inteiros *a* e *b*, ou seja, é uma relação *simétrica*.
(3) Se $a \equiv b(\mathrm{mod}\, n)$ e $b \equiv c(\mathrm{mod}\, n)$, então $a \equiv c(\mathrm{mod}\, n)$ para quaisquer inteiros *a*, *b* e *c*, ou seja, é uma relação *transitiva*.

Como as três propriedades são atendidas, a relação de congruência é uma relação de equivalência. A demonstração das propriedades é imediata, haja vista que, conforme dito anteriormente, dois números são congruentes quando deixam o mesmo resto na divisão por *n*. Além disso, podemos verificar que a relação de congruência é preservada por somas e por produtos.

Proposição 3.4.1 Se $a \equiv b(\mathrm{mod}\, n)$ e $c \equiv d(\mathrm{mod}\, n)$, então $a+c \equiv b+d(\mathrm{mod}\, n)$ e $ac \equiv bd(\mathrm{mod}\, n)$.

Demonstração. Se $a \equiv b(\mathrm{mod}\, n)$, então $n \mid a-b$. Do mesmo modo, como $c \equiv d(\mathrm{mod}\, n)$, então $n \mid c-d$. Daí, *n* é múltiplo de $a-b$ e de $c-d$. Como a soma de múltiplos de *n* é também um múltiplo de *n*, então $n \mid a-b+c-d$, ou seja, $a+c \equiv b+d(\mathrm{mod}\, n)$.
Se $n \mid a-b$, ocorre também que $n \mid (a-b)\cdot c$ e, se $n \mid c-d$, então $n \mid (c-d)\cdot b$. Utilizando o mesmo argumento anterior, temos que $n \mid (a-b)\cdot c+(c-d)b$. Daí, temos que $n \mid ac-bc+bc-bd \Rightarrow n \mid ac-bd$. Logo, $ac \equiv bd(\mathrm{mod}\, n)$.

Várias outras propriedades importantes decorrem da definição de congruência. Os tópicos a seguir, que também fazem uso das congruências, apresentam teoremas que nos permitem realizar testes de primalidade que nos serão bastante úteis no trabalho com codificação e decodificação.

3.5 Pequeno Teorema de Fermat

Aproveitando o que foi visto no capítulo anterior e nos tópicos já apresentados sobre a teoria dos números, vamos provar o Teorema de Fermat, que nos permite encontrar números primos que podem nos ajudar na implementação do método de criptografia conhecido como RSA. No RSA é preciso escolher números primos bem grandes. Sugere-se a escolha de primos com 100 dígitos ou mais.

Para demonstrar o Pequeno Teorema de Fermat, necessitamos da demonstração e compreensão do lema a seguir:

Lema 3.5.1 Seja *p* um número primo. Os números $\binom{p}{i}$, onde $0 < i < p$, são todos divisíveis por *p*.

Demonstração. O resultado vale trivialmente para $i = 1$. Portanto, podemos supor que $0 < i < p$. Nesse caso, $i!\mid p\cdot(p-1)\cdot\ldots\cdot(p-i+1)$. Como $(i!, p) = 1$, decorre que $i!\mid (p-1)\cdot\ldots\cdot(p-i+1)$ e o resultado segue, pois $\binom{p}{i} = p\cdot\dfrac{(p-1)\cdot\ldots\cdot(p-i+1)}{i!}$.

Teorema 3.5.1 (Pequeno Teorema de Fermat) Dado um número primo *p*, tem-se que *p* divide o número $a^p - a$, ou seja, que $a^p \equiv a(\mathrm{mod}\, p)$, para todo $a \in \mathrm{Z}$.

Demonstração. Basta mostrar o resultado para $a \geq 0$. Vamos provar por indução sobre *a*. O resultado vale para $a = 0$, pois $p \mid 0$. Admitindo que o resultado vale para *a*, vamos provar sua validade para $a+1$. Pela fórmula do Binômio de Newton, temos:

$$(a+1)^p - (a+1) = \binom{p}{0} \cdot a^{p-0} \cdot 1^0 + \binom{p}{1} \cdot a^{p-1} \cdot 1^1 + \ldots + \binom{p}{p} \cdot a^{p-p} \cdot 1^p - a - 1$$

Como $p \mid a^p - a$ (por hipótese de indução) e $p \mid \binom{p}{i}$, para $0 \leq i \leq p$, ou seja, o segundo membro da equação é divisível por *p* (Lema 3.5.1), temos que $p \mid (a+1)^p - (a+1)$. Logo, vale para todo $a \geq 0$. A prova para $a < 0$ é realizada de modo análogo.

□

Corolário 3.5.1 Se *p* é um número primo e *a* é um inteiro não divisível por *p*, então *p* divide $a^{p-1} - 1$, ou seja, $a^{p-1} \equiv 1 (\mathrm{mod}\, p)$.

Demonstração. Como, pelo pequeno Teorema de Fermat, $p \mid a^p - a$, então $p \mid a \cdot (a^{p-1} - 1)$. Se *a* é um inteiro não divisível por *p*, então $mdc(a, p) = 1$. Daí, segue-se, imediatamente, que $p \mid a^{p-1} - 1$.

Para determinar o resto da divisão de 2^{120} por 7, por exemplo, podemos utilizar o corolário do Pequeno Teorema de Fermat do seguinte modo: como 2 não é divisível por 7, temos $2^{7-1} \equiv 1 (\mathrm{mod}\, 7) \Rightarrow 2^6 \equiv 1 (\mathrm{mod}\, 7)$. Elevando ambos os membros a 20, teremos $2^{120} \equiv 1 (\mathrm{mod}\, 7)$, ou seja, a potência deixa resto 1 quando dividida por 7.

3.6 Teorema de Euler

Do mesmo modo que o Pequeno Teorema de Fermat, o Teorema de Euler também nos fornece um teste de primalidade. Para compreendê-lo, precisamos conhecer a função ϕ (phi) de Euler. A função associa, para cada inteiro positivo *n*, o valor $\phi(n) = \#\{k \in Z; 0 < k < n \text{ e } mdc(k, n) = 1\}$, ou seja, a função associa a cada número inteiro positivo *n* a quantidade de inteiros positivos relativamente primos com *n*.

Teorema 3.6.1 Para *p* primo e *a* um inteiro positivo, temos $\phi(p^a) = p^a - p^{a-1}$.

Demonstração. Pela definição de $\phi(n)$, sabemos que $\phi(p^a)$ é o número de inteiros positivos não superiores a p^a e relativamente primos com p^a. Mas, os únicos não primos com p^a e menores do que ou iguais a p^a são aqueles divisíveis por *p*. Como os múltiplos de *p* não superiores a p^a são, em número, p^{a-1}, o resultado segue.

Para qualquer primo *n*, todos os inteiros positivos menores que *n* são relativamente primos com *n*, ou seja, $\phi(n) = n - 1$. Além disso, podemos provar também que a função $\phi(n)$ é multiplicativa, o que faremos a seguir.

Teorema 3.6.2 A função ϕ de Euler é multiplicativa, isto é, $\phi(mn) = \phi(m) \cdot \phi(n)$ para $mdc(m,n) = 1$.

Demonstração. Vamos dispor os números de 1 até *mn* da seguinte forma:

1	$m + 1$	$2m + 1$	...	$(n - 1)\, m + 1$
2	$m + 2$	$2m + 2$	...	$(n - 1)\, m + 2$
3	$m + 3$	$2m + 3$	...	$(n - 1)\, m + 3$
...	...	...	...	...
m	$2m$	$3m$	...	nm

Se na linha *r*, onde estão os termos *r, m + r, 2m + r, ..., (n – 1)m + r*, tivermos $mdc(m,r) = d > 1$, então nenhum termo nesta linha será primo com *mn*, uma vez que estes termos, sendo da forma $km + r$, $0 \le k \le n - 1$, são todos divisíveis por *d*, de modo que $d = mdc(m,r)$. Logo, para encontrarmos os inteiros desta tabela que são primos com *mn*, devemos olhar na linha *r* somente se $mdc(m,r) = 1$. Portanto, temos $\phi(m)$ linhas, onde todos os elementos são primos com *m*. A seguir, devemos procurar em cada uma dessas $\phi(m)$ linhas quantos elementos são primos com *n*, uma vez que todos são primos com *m*. Como $mdc(m,n) = 1$, os elementos *r, m + r, 2m + r, ..., (n – 1)m + r* formam um sistema completo de resíduos módulo *n*. Logo, cada uma dessas linhas possui $\phi(n)$ elementos primos com *n* e, portanto, como eles, são primos com *m* e também são primos com *mn*. Isto nos garante que $\phi(mn) = \phi(m) \cdot \phi(n)$. □

Então, para determinar $\phi(100)$, por exemplo, faremos: $\phi(100)=\phi(2^2\cdot 5^2)=\phi(2^2)\cdot\phi(5^2)$. Logo, teremos: $\phi(2^2)\cdot\phi(5^2)=(2^2-2^{2-1})\cdot(5^2-5^{2-1})$. Portanto, $\phi(100)=\phi(2^2)\cdot\phi(5^2)=2\cdot 20=40$.

Antes de demonstrar o principal teorema do tópico, apresentaremos um outro teorema que ajudará na sua compreensão.

Teorema 3.6.3 Seja *a* um inteiro positivo tal que $mdc(a,n)=1$. Se $r_1.,r_2,..,r_{\phi(n)}$ é um sistema reduzido de resíduos módulo *n*, então $ar_1,ar_2,...,ar_{\phi(n)}$ também o é.

Uma vez tendo o resultado anterior em mãos e já que entendemos como funciona a função de Euler, vamos ao teorema que dá nome ao tópico.

Teorema 3.6.4 (Teorema de Euler) Sejam *n* e *a* inteiros relativamente primos, então $a^{\phi(n)}\equiv 1(\mathrm{mod}\, n)$.

Demonstração. De acordo com o teorema 3.6.3, os elementos $ar_1,ar_2,...,ar_{\phi(n)}$ constituem um sistema reduzido de resíduos módulo *n* se $mdc(a,n)=1$ e $r_1.,r_2,..,r_{\phi(n)}$ for um sistema reduzido de resíduos módulo *n*. Isto significa que ar_i é congruente a exatamente um dos r_j, $1\le j\le\phi(n)$ e, portanto, o produto dos ar_i deve ser congruente ao produto dos r_j módulo *n*, isto é, $ar_1.\cdot ar_2\cdot..\cdot ar_{\phi(n)}\equiv r_1\cdot....\cdot r_{\phi(n)}(\mathrm{mod}\, n)$, ou seja, $a^{\phi(n)}\cdot r_1.\cdot r_2\cdot..\cdot r_{\phi(n)}\equiv r_1\cdot....\cdot r_{\phi(n)}(\mathrm{mod}\, n)$. Como $\left(\prod_{i=1}^{\phi(n)} r_i,n\right)=1$, podemos cancelar o produto $r_1.\cdot r_2\cdot..\cdot r_{\phi(n)}$ em ambos os lados, o que nos leva a concluir que $a^{\phi(n)}\equiv 1(\mathrm{mod}\, n)$.

□

Uma vez que conseguimos compreender os tópicos da Teoria dos Números apresentados nesse capítulo, iremos, agora, adentrar no campo da criptografia que encontra-se no texto dividida em dois capítulos – III e IV, onde tentaremos compreender, respectivamente, como se deu seu desenvolvimento até os dias de hoje e de que modo os números primos contribuem na codificação e decodificação de mensagens, senhas (sobretudo as utilizadas na *internet*) etc.

4 CAPÍTULO III – O ESTUDO DA CRIPTOGRAFIA: O QUE É, SURGIMENTO, DESENVOLVIMENTO E RSA NA INTERNET

Há muito tempo os indivíduos procuram meios de comunicar-se. A comunicação é um processo que envolve troca de informações, ideias, ensinamentos, pensamentos, mensagens, entre outros. Acredita-se que a necessidade de comunicação se deve ao surgimento da vida em sociedade. Por conta disso, em um determinado momento do passado, o homem de comunicação rudimentar acabou criando uma forma primitiva e simples de linguagem que, com o tempo, foi adquirindo formas mais claras e evoluídas, facilitando o intercâmbio e a troca de informações não só entre os povos de uma mesma tribo, se estendendo a povos de tribos vizinhas.

O surgimento da escrita, evolução da comunicação que para muitos estudiosos marca o início da história, aparece frente a uma necessidade do desenvolvimento da economia e da sociedade há cerca de 8600 anos atrás. A troca de mensagens escritas, nesse período, ajudou a democratizar as informações e promoveu a possibilidade de torná-las acessíveis a um número cada vez maior de pessoas.

É importante ressaltar que desde o advento da escrita surgiu uma necessidade crescente de enviar mensagens secretas, ou seja, mensagens que só o emissor e o receptor possam decifrar. Tal necessidade fez surgir a criptografia, que tem por base a escrita de mensagens a partir de códigos secretos.

Nesse sentido, o presente capítulo procura explicar o que é a criptografia e para que pode ser utilizada, além de apresentar um histórico sobre o advento da utilização de códigos e os passos de seu desenvolvimento até os dias atuais. O capítulo também apresenta um tópico que se destina ao histórico da criptografia RSA desde sua idealização até sua implementação, que é indispensável, dentre outras coisas, às transações do comércio eletrônico atual.

4.1 O que é Criptografia?

Criptografia (do grego *cryptós* - “secreto, escondido, oculto”; e *gráphein*- “escrita”). É a área da Criptologia (ciência da encriptação) que se ocupa da escrita em códigos e das técnicas de transformação dessa escrita. Nesse sentido, a criptografia destina-se a estudar os modos como uma mensagem pode ser desfigurada de sua forma original a fim de

obter outra aparência, alheia aos curiosos a quem não se destina, de modo que apenas possa ser conhecida por aqueles que possuem o código necessário à sua decodificação.

Com uma utilização que remota aos tempos antigos, a criptografia tem hoje importante papel na comunicação global. Trataremos no próximo tópico do histórico sobre o surgimento e o desenvolvimento da criptografia, desde sua utilização com fins militares até o modo como é feita nos dias atuais.

4.2 Surgimento e desenvolvimento

A criptografia foi criada com o intuito de elaborar e reproduzir mensagens utilizando uma chave ou código, possibilitando que uma mensagem interceptada por um leitor que não é o devido destinatário apresente-se, pelo menos a princípio, como algo ilegível. Consideramos que a mensagem é "a princípio" ilegível, pois o desenvolvimento da criptografia é imediatamente acompanhado do desenvolvimento da criptoanálise, ciência que faz justamente o oposto: decifra a lógica empregada, ou seja, descobre a cifra que restringe a mensagem criptografada ao emissor e ao receptor detentores de seu código. Desse modo, o interceptador das mensagens pode lê-las mesmo sem saber de imediato qual é a chave ou código empregados na sua inscrição.

Além disso, é natural que uma mensagem cifrada possua duas partes: a codificação e a decodificação. Decodificar não é o mesmo que decifrar. A pessoa responsável por decodificar uma mensagem é o seu receptor ou destinatário. Isso quer dizer que quem decodifica uma mensagem possui o código que permite sua leitura. Já para decifrar, a pessoa não necessariamente tem que saber o código, basta que ela tenha perícia suficiente para "quebrá-lo".

Um dos primeiros marcos históricos quanto à decifração de códigos foi a descoberta do significado dos *hieróglifos* (cada um dos sinais de escrita usado pelas antigas civilizações). A decifração dos hieróglifos se deu a partir da descoberta da chamada Pedra de Roseta. A esse respeito, conta-se que mais ou menos em 1799, durante a campanha de Napoleão Bonaparte no Egito, um soldado francês encontrou uma pedra com inscrições antigas perto da cidade de Roseta. A pedra continha três parágrafos de inscrições e, posteriormente, foi descoberto que cada parágrafo estava escrito em um idioma diferente. O método utilizado pelo decifrador, nesse caso, foi a contagem e frequência de caracteres.

Com o tempo, muito mais do que decifrar mensagens antigas, ocorreu às pessoas, em especial aos grandes líderes, que a utilização de códigos secretos com fins militares poderia ser de grande valor.

César, por exemplo, utilizou o que se acredita ser o mais simples dos códigos. A "Cifra de César", como passou a ser chamado, consistia na substituição de letras. Considere, por exemplo, a mensagem: TRABALHO CONCLUIDO COM SUCESSO. Admitindo o alfabeto cíclico e substituindo uma letra pela seguinte, teremos: USBCBMIP DPODMVJEP DPN TVDFTTP. Assim, usando sua cifra, César comunicou-se com sua legião de combate por toda a Europa.

O exército espartano, há mais de 2500 anos atrás, também fez uso de um método de codificação. No caso espartano, o método consistia em escrever mensagens enrolando um pergaminho em um cilindro (chamado de cítala). Quando desenrolado, o pergaminho não fazia sentido a quem não possuísse um cilindro com as mesmas características daquele utilizado para escrever a mensagem inicial.

Apesar da simplicidade dos códigos apresentados, sua eficiência não pode deixar de ser considerada. Alguém que desejasse interceptar essas comunicações deveria descobrir de que modo essas mensagens foram escritas para poder ter acesso ao seu conteúdo.

Durante muito tempo, códigos desse tipo foram utilizados e não tivemos nenhuma novidade no campo da criptografia, até que, por volta da década de 1840, Edgard Allan Poe (autor, poeta, editor e crítico literário) teve seu interesse despertado pela resolução de enigmas. O poeta publicou suas habilidades no jornal da Filadélfia, Alexander´s Weekly, solicitando códigos para resolver e, tempos depois, escreveu um ensaio sobre métodos de criptografia, que se tornou útil aos criptoanalistas britânicos na quebra dos códigos alemães durante a Primeira Guerra Mundial.

Durante a guerra, vários códigos alemães foram quebrados e seus navios interceptados. No entanto, a contribuição mais importante dada pelos criptoanalistas britânicos nesse período foi a decodificação do telegrama de Zimmermann. No telegrama, enviado ao embaixador alemão no México, Heinrich Von Eckardt, havia instruções para que o embaixador procurasse uma aproximação com governo mexicano a fim de propor uma aliança militar contra os Estados unidos. O telegrama foi interceptado pela Inglaterra e enviado ao governo norte-americano, o que não só apressou como foi um marco decisivo para a entrada dos Estados Unidos na guerra.

Tempos depois, já com o fim da Primeira Grande Guerra, destacamos um dos métodos de codificação mais famosos, patenteado em 1918 pelo engenheiro elétrico e inventor alemão Arthur Scherbius: a máquina Enigma (Figura 9). A máquina, semelhante a uma máquina de escrever não atraiu de início tanta atenção do público e nem vingou comercialmente, chegando a ser oferecida para a Marinha alemã. A Marinha, por sua vez, resolveu indicar o Escritório de Relações Exteriores, para quem a máquina foi apresentada com fins diplomáticos, mas, mesmo assim, ainda não havia interesse. Nesse período, a patente da máquina passou por vários inventores que travavam batalhas na reestruturação de máquinas parecidas.

Figura 9

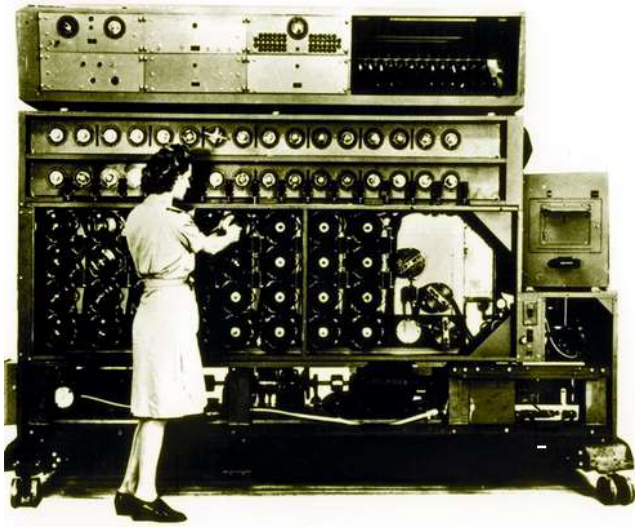
Fonte: Disponível em: <http://supercurioso.com/la-misteriosa-maquina-enigma/>. Acesso em: 19/03/15.

Em 1926 a máquina foi finalmente adotada pela Marinha alemã, que comprou alguns exemplares e adaptou-os, até que em 1928 o exército elaborou a sua própria versão. A partir desse momento, o uso da máquina Enigma estendeu-se a toda a organização militar alemã. A versão alemã da máquina foi amplamente utilizada durante a Segunda Guerra Mundial e passou a transmitir praticamente todas as comunicações de rádio, tal como as comunicações telegráficas e até mesmo os boletins meteorológicos.

Após o estopim da guerra, um grupo de criptógrafos britânicos (muitos deles mestres em xadrez e matemáticos, como William Gordon Welchman, Max Newman e Alan Turing, fundador conceitual da computação moderna), trabalhando em Bletchley Park, conseguiu quebrar os códigos da máquina Enigma e decifrar as mensagens secretas dos nazistas.

As máquinas alemãs usavam uma cifra de chave única, ou seja, um algoritmo onde o texto é combinado com uma chave aleatória. Já os ingleses, liderados por Turing, após utilizar uma série de técnicas que se mostraram a princípio infrutíferas, decifraram a chave e criaram o primeiro computador digital programável. O computador Colossus (Figura 10), como foi chamado, quebrou os códigos da Enigma e, segundo alguns especialistas, adiantou o fim da guerra em mais ou menos dois anos.

Figura 10



Fonte: Disponível em: <http://brasileiros.com.br/2012/06/o-homem-da-tecnologia/>. Acesso em: 19/03/15.

É importante lembrar que toda a história da criptografia até o ponto destacado dependeu de uma comunicação entre emissor e receptor, que deveriam encontrar-se para definir a cifra utilizada na conversa. Sautoy (2007, p. 242) destaca a esse respeito que:

> antes de 1977, quem quisesse enviar uma mensagem secreta depararia com um problema essencial. Antes que o comunicado fosse transmitido, o emissor e o receptor teriam de se encontrar para decidir qual cifra – o método de codificação – usariam. Os generais espartanos, por exemplo, precisavam concordar sobre as dimensões da cítala. Mesmo com a máquina Enigma, produzida em série, Berlin tinha que enviar agentes para fornecer aos capitais dos barcos U e aos comandantes dos tanques os livros que descreviam as configurações da máquina para codificar as mensagens de cada dia. Naturalmente, se um inimigo pusesse as mãos no livro de códigos, o jogo terminava. (p. 242)

Pensando nisso, as sociedades passaram a desenvolver métodos cada vez mais sofisticados para codificar suas mensagens. A decifração de mensagens com o uso de uma máquina equivalente a um computador mostrou-se eficiente no período da guerra e a utilização de computadores a partir desse ponto foi cada vez mais necessária.

Os computadores trouxeram ainda mais desafios à criptografia. A comunicação entre eles, proporcionada pela *internet*, é um exemplo. Assim, encontramo-nos em um momento onde as pessoas distantes fisicamente precisam se comunicar e o computador é uma ferramenta bastante útil para propiciar tal intento. Desse modo, para que a comunicação entre computadores se dê de forma efetiva e eficiente, é preciso que a mensagem navegue pela rede sem ser decifrada por qualquer um.

Usar uma das técnicas que demanda da entrega de chaves para os usuários com a finalidade de fazer negócios pela *internet* seria deveras arriscado, mesmo porque quanto maior a distância entre as máquinas, mais complicada é a logística envolvida no trabalho. As chaves enviadas aos usuários da maneira convencional como era feita, a partir dos correios ou de um mensageiro, poderiam ser interceptadas e a espera para fazer esse tipo de transação, poderia tornar-se bem prolongada.

Pensando nisso, fez-se necessária a criação de códigos difíceis de decifrar mesmo com a ajuda do computador. A criptografia avançou mais um passo até a criação de um sistema, teoricamente seguro, denominado sistema de códigos de *chave pública*. Sabe-se que:

> a criptografia de chave pública foi proposta inicialmente em 1976, em um artigo seminal escrito por dois matemáticos da Universidade de Stanford, na Califórnia, Whit Diffie e Martin Hellman. A dupla desencadeou o surgimento de uma contracultura do mundo criptográfico, que passaria a desafiar o monopólio das agências governamentais sobre a criptografia. [...] Os dois apoiavam a ideia de que a criptografia não deveria ser um assunto discutido entre as portas fechadas no governo, mas que suas ideias deviam se tornar públicas para beneficiar a todos. (SAUTOY, 2007, p. 243)

A criptografia de chave pública é um método que utiliza um par de chaves: uma pública e uma chave privada. A chave pública, como diz o nome, é distribuída livremente, enquanto a chave privada é de conhecimento bastante restrito. Nesse sentido:

> o sistema de criptografia de chave pública é como uma porta com duas chaves diferentes: a chave A tranca a porta, mas uma chave diferente, B, a destranca. Então, não é mais necessário manter qualquer confidencialidade em relação à chave A. Distribuir cópias dela não compromete a segurança. (SAUTOY, 2007, p. 243)

Na criptografia de chave pública uma mensagem cifrada só pode ser decifrada por sua chave privada correspondente. Sautoy (2007, p. 243) exemplifica:

> a empresa pode distribuir livremente a chave A para qualquer visitante da página que queira enviar uma mensagem segura, como o número de um cartão de crédito. Embora todos usem a mesma chave para codificar seus dados – trancando a porta e

> protegendo o segredo –, ninguém consegue ler as mensagens codificadas dos demais. Na verdade, uma vez codificados os dados, os clientes não conseguem mais lê-los, mesmo que sejam os seus próprios. Somente a empresa que administra a página virtual possui a chave B, usada para destrancar a porta e ler os números dos cartões de crédito.

No fim da década de 1970 a ideia estava lançada. O problema era a construção de um algoritmo de chave pública que permitisse a total confidencialidade de quem deposita os seus dados na grande rede. Como a criptografia vinha atraindo a atenção de um número cada vez maior de pessoas, a solução para o problema não tardou a chegar.

No próximo tópico, apresentaremos a solução para o problema da construção de um algoritmo de chave pública, bem como sua popularização na *internet*. A criptografia RSA, como é chamado o método mais famoso que utiliza de uma chave pública, será apresentada desde sua idealização.

4.3 RSA na *internet*

Após ler o artigo de Diffie e Hellman sobre a criptografia de chave pública, Ron Rivest (integrante do Departamento de Ciências da Computação do MIT – Instituto de Tecnologia de Massachusetts) interessou-se bastante pelo assunto e percebeu muito precocemente que isso teria várias implicações práticas no mundo real.

Já que a criptografia consistia na elaboração e utilização de códigos difíceis de decifrar, Rivest acreditava que quanto mais complicado fosse o problema empregado na elaboração desses códigos, mais laborioso seria o trabalho dos criptoanalistas na sua quebra. Além disso, em uma análise mais profunda, dependendo do problema, sua resolução poderia ser impossível.

Assim, um de seus desafios principais, segundo Rivest, era distinguir os problemas fáceis dos complicados, ainda mais se a solução dessas questões dependesse de máquinas como os computadores, que podem ser exímios solucionadores de problemas. Assim, ao decidir dedicar seus esforços à construção de um sistema de criptografia de chave pública, Rivest percebeu que para dificultar a quebra dos códigos utilizados para codificar as mensagens a serem enviadas, deveria levar em conta problemas com soluções difíceis de calcular, até mesmo para os computadores. Desse modo:

> Rivest se dedicou a criar um sistema de criptografia de chave pública explorando todo tipo de problema que os computadores sabidamente levariam um longo tempo

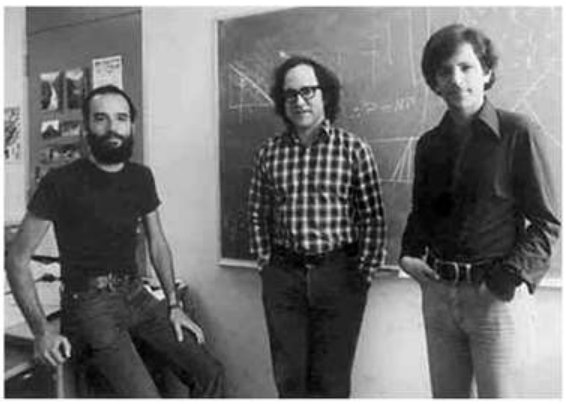
> para resolver. Ele também precisava de alguém para testar suas ideias. O MIT já começava a romper os moldes das universidades tradicionais, afrouxando as fronteiras entre os departamentos na esperança de incentivar as interações disciplinares. Rivest, um cientista da computação, trabalhava no mesmo pavimento que integrantes do departamento de matemática. Em escritórios próximos havia dois matemáticos, Leonard Adleman e Adi Shamir. (SAUTOY, 2007, p. 245)

Contando com o apoio dos colegas Adi Shamir (matemático, mestre e doutor em Ciências da Computação) e Leonard Adleman (matemático, professor de Ciências da Computação e Biologia Molecular), Rivest acabou adentrando no mundo da matemática onde explorou diversos problemas difíceis. Sautoy (2007, p. 245-246) destaca:

> enquanto exploravam os diversos problemas matemáticos "difíceis", seus sistemas criptográficos embrionários começaram a usar mais ideias da teoria dos números. [...] Eles vinham pensando há algum tempo no difícil problema de fatorar números. Não havia boas propostas de programas que conseguissem decompor números em seus blocos de construção primos. O problema parecia ser do tipo que buscavam.

Ao se deparar com o problema da fatoração, o trio acreditava ter encontrado a chave que tornaria complicada a decifração das mensagens codificadas. Assim, Rivest escreveu um manuscrito sobre a descoberta do trio do MIT (Figura 11) e, depois de intensas discussões sobre a divisão dos créditos da ideia (coisa que Adleman atribuía principalmente a Rivest), o sistema ficou conhecido como criptografia RSA.

Figura 11



Fonte: Disponível em: <http://viterbi.usc.edu/news/news/2011/len-adleman-and.htm>. Acesso em: 25/03/15.

Posteriormente à escrita do projeto, Rivest se perguntava se seria realmente tão difícil a fatoração de números. Nesse momento, Martin Gardner, grande colaborador da divulgação científica e matemática, colunista da *Scientific American*, intrigado com a proposta

de Rivest sobre a fatoração e o manuscrito que propunha uma criptografia de chave pública, perguntou se podia publicar um artigo sobre o assunto.

No artigo publicado, Gardner garantia que o trio estava disposto a enviar sua pesquisa a qualquer pessoa que mandasse a eles um envelope selado. Qual foi a surpresa dos estudiosos quando vários envelopes chegaram pelos correios? Nesse momento, eles começaram a acreditar com mais força que seu projeto tinha ares de grandeza que sua imaginação ainda não tinha alcançado. A essa altura,

> o trio começava a ouvir que ficaria rico. Corriam os anos de 1970, ninguém ainda pensava no comércio eletrônico, mas as pessoas entenderam a força daquelas ideias. [...] Ao que tudo indicava, a mesma ideia fora proposta a portas fechadas no mundo da inteligência. Mas, as agências de segurança não tinham muita certeza sobre se deveriam colocar as vidas de agentes nas mãos de uns poucos matemáticos que diziam que era difícil decifrar números. Ansgar Heuser, do BSI, a Agência de Segurança Nacional alemã, lembra que nos anos 1980 considerou-se pôr em prática o RSA. (SAUTOY, 2007, p. 247-248)

O trio precisava se convencer, além de convencer o mundo, que o problema da fatoração era, de fato, sólido para os seus propósitos. Dessa maneira, no ato do lançamento do mencionado artigo, estudaram sobre o grau de dificuldade de decompor um número em seus constituintes primos e chegaram à conclusão que deveriam utilizar números com mais de 100 algarismos. Nesse momento, resolveram lançar o desafio denominado RSA129. Assim, ao encontrarem-se:

> com tantos números para verificar, Rivest, Shamir e Adleman estavam suficientemente confiantes para lançar um desafio: decifrar um número com 129 algarismos que haviam criado a partir de dois primos. [...] No artigo, estimaram que seriam necessários cerca de 40 quadrilhões de anos para decifrar o RSA129. Logo perceberam que haviam cometido um pequeno deslize aritmético nessa estimativa. Ainda assim, com as técnicas de fatoração de números disponíveis na época, o processo levaria milhares de anos. (SAUTOY, 2007, p. 253)

O desafio foi lançado e milhares de matemáticos se puseram a trabalhar na descoberta do RSA129. Levaram 17 anos desde o desafio para que o número fosse decifrado. Mesmo que o período para a fatoração do RSA129 não tenha chegado nem perto da estimativa do trio, Sautoy (2007, p. 254) deixa claro que 17 anos era um período longo para que, por exemplo, a data de validade de um cartão de crédito codificado com o RSA129 já tivesse expirado. Além disso, se o RSA129 foi decifrado, restava ao trio aumentar o tamanho dos primos e foi o que fizeram. Assim,

> no final dos anos 1990, Rivest, Shamir e Adleman fizeram uma nova série de desafios. No final de 2002, o menor deles ainda não decifrado era um número com 160 algarismos. [...] Rivest se desfez dos primos que usou para gerar os números desses desafios, portanto ninguém conhece realmente as respostas até que os números sejam decifrados. (SAUTOY, 2007, p. 256-257)

Em consequência do sucesso obtido com os desafios, o trio pôde atestar que o problema atendia às necessidades do algoritmo criado por eles. Dessa forma, mesmo com reservas iniciais, a criptografia RSA foi posta em prática e, desde o advento da *internet*, é utilizada na maior parte das transações que emitem dados para a grande rede. O sistema mostrou-se bastante seguro, pois estamos lidando com a transmissão de dados importantes e confiando informações a um algoritmo de criptografia. Uma vez que o algoritmo seja decifrado, qualquer um pode ter acesso às nossas informações confidenciais.

Nesse sentido, no capítulo seguinte, voltaremos ao assunto da criptografia RSA, explicando-a melhor, apresentando exemplos práticos e esclarecendo a relação íntima que os números primos e o que os tópicos de teoria dos números apresentados nos capítulos I e II têm a ver com esse tema.

5 CAPÍTULO IV – CRIPTOGRAFIA RSA: DA TEORIA À PRÁTICA

Já faz um tempo que a importância dos números primos deixou de ser puramente acadêmica. A descoberta de primos que constituem números grandes – passatempo apreciado por muitos matemáticos – encontra-se hoje, de acordo com Sautoy (2007, p. 241), "no centro da moderna decifração de códigos". Vale lembrar que há:

> mais de dois mil anos atrás, os gregos provaram que todos os números podem ser expressos como um produto de primos. Desde então, os matemáticos têm tido dificuldade em encontrar um método rápido e eficiente para descobrir quais foram os números primos usados para gerar outros números. Ainda não temos um correspondente matemático da espectroscopia, que diz aos químicos quais elementos da tabela periódica formam um composto químico. A descoberta de um homólogo matemático conseguiria decifrar os constituintes primos e renderia a seu criador muito mais que o mérito acadêmico. (SAUTOY, 2007, p. 240-241)

A criptografia, em uma análise prática, é uma aplicação que leva a teoria que envolve a codificação para a realidade palpável do mundo em que vivemos. Sautoy (2007, p. 241) nos lembra que: "Para impedir que informações importantes caíssem em mãos erradas, nossos ancestrais inventaram maneiras cada vez mais perspicazes de dissimular o conteúdo de uma mensagem".

Assim, para criptografar, codificar ou tornar uma mensagem incompreensível é preciso definir um protocolo de encriptação, denominado chave. As chaves utilizadas para encriptar mensagens podem ser simétricas ou assimétricas, dependendo da criptografia empregada. A criptografia simétrica utiliza apenas uma chave para codificar e decodificar uma mensagem, enquanto a assimétrica (que tem a RSA como exemplo) utiliza duas chaves: uma pública (distribuída livremente) e uma chave privada.

Para poder adentrar no mundo da criptografia RSA de duas chaves e compreender a utilização que a *internet* faz desse método, precisamos nos valer dos tópicos da teoria dos números, apresentados no capítulo II e, é claro, de nossos arquivos sobre os números primos (capítulo I).

Logo, cada tópico do capítulo procurará explicar o funcionamento do algoritmo que sustenta as bases da RSA, apresentando uma pequena introdução ao método, o algoritmo propriamente dito que é usado para a encriptação e desencriptação de mensagens, exemplos básicos à aplicação do mencionado algoritmo, métodos para a fatoração de inteiros, além de explicar o que leva à segurança e funcionamento do método.

5.1 Introdução ao método RSA

Destacando-se por ser um dos métodos mais utilizados na atualidade, principalmente no comércio eletrônico, a criptografia RSA (exemplo de criptografia assimétrica) conta com duas chaves: uma pública e uma privada. Seu funcionamento baseia-se na facilidade de emissão da chave pública (ou chave de encriptação) e na impossibilidade de quebrar ou descobrir a chave privada de decodificação ou desencriptação. Tal dificuldade é ocasionada pela inexistência de algoritmos com baixo custo computacional e suficientemente eficientes para a fatoração de números com uma grande quantidade de algarismos (base para a escrita RSA).

Quanto maior o número empregado no algoritmo da codificação, maiores serão os primos empregados na sua constituição. A esse respeito, Sautoy (2007, p. 261) acredita que a ideia de Euclides sobre a infinidade dos primos atende à demanda da *internet* por números cada vez maiores e o fato dos primos nunca esgotarem "adquiriu subitamente um significado comercial inquestionável". Além disso, o autor destaca que:

> felizmente, a natureza foi caridosa com o mundo do comércio eletrônico. O teorema dos números primos de Gauss determina que a quantidade de primos com 60 algarismos é de aproximadamente 10^{60} dividido pelo logaritmo de 10^{60}. Ou seja, existem suficientes primos com 60 algarismos para que cada átomo da Terra tenha seu próprio par de primos. (SAUTOY, 2007, p. 261)

Assim, fica claro que sempre teremos primos (com 60 algarismos como menciona o autor ou mais) para alimentar o algoritmo RSA. Portanto, de maneira prática, o algoritmo que sustenta as ideias de Rivest, Shamir e Adleman funciona como uma receita de bolo e segue um conjunto de instruções que envolvem números primos muito grandes e algumas ideias básicas de teoria dos números. Embora o comércio eletrônico considere no algoritmo a utilização de números que tenham mais de 100 algarismos, vamos expor adiante exemplos mais didáticos para a compreensão do leitor.

5.2 O Algoritmo RSA

Devemos supor, inicialmente, que as mensagens consideradas para fins de codificação tenham apenas letras que formam as palavras (sem números) e espaços entre as

palavras. Assim, o passo (0), que não é colocado a seguir, consiste em converter as letras e espaços da mensagem a ser criptografada em números.

Para poder proceder com o passo (0), devemos escolher uma matriz de substituição para a mencionada troca de letras por números. Utilizaremos nos exemplos do tópico seguinte os dados da Tabela 2, onde cada letra e cada espaço correspondem a um número de dois algarismos. Observe que estamos considerando números de dois algarismos para evitar ambiguidades que poderiam ocorrer se começássemos o trabalho com números de apenas um algarismo. Se assim fosse, não saberíamos a que se referia 13, por exemplo, se a AC ou a M.

Tabela 2 – Matriz de substituição para o algoritmo RSA

A	B	C	D	E	F	G	H	I	J	K	L	M
10	11	12	13	14	15	16	17	18	19	20	21	22

N	O	P	Q	R	S	T	U	V	W	X	Y	Z	-
23	24	25	26	27	28	29	30	31	32	33	34	35	99

Fonte: Elaborado pela autora

Portanto, uma vez compreendido que devemos iniciar os trabalhos com a substituição de letras por números, ou seja, com uma espécie de "pré-codificação", vamos ao passo a passo do algoritmo.

(1) Escolhem-se primos p e q (geralmente com mais de 100 algarismos);

(2) Determinam-se $n = pq$ e $\phi(n) = (p-1)\cdot(q-1)$;

(3) Escolhe-se um inteiro positivo e de modo que $mdc(e, \phi(n)) = 1$;

(4) Chamamos o par (n, e) de *chave de codificação* do sistema RSA. Agora, a mensagem deve ser codificada em blocos b. Denotaremos cada bloco codificado por $C(b)$, de modo que $C(b)$ é o resto da divisão de b^e por n, ou seja, $C(b) \equiv b^e \pmod{n}$;

(5) Escolhe-se um inteiro positivo d, de modo que $d \cdot e \equiv 1 \pmod{\phi(n)}$;

(6) Chamamos o par (n,d) de *chave de decodificação* do sistema RSA. Agora, a mensagem deve ser decodificada em blocos *a*. Denotaremos cada bloco decodificado por $D(a)$, de modo que $D(a)$ é o resto da divisão de a^d por *n*, ou seja, $D(a) \equiv a^d \pmod{n}$.

Uma vez que a mensagem foi transformada em um grande número, em (0), esse número deve ser quebrado em blocos, para que a codificação possa ser feita. Os blocos, apesar de poder ter tamanhos e maneiras distintas de escolha, devem conter números menores que *n* e não devem começar com 0, pois isso dificultaria o trabalho de decodificação. Como distinguir os blocos 31 e 031? Além disso, a codificação em blocos é necessária, pois os mesmos não correspondem a nenhuma unidade linguística e isso torna inviável a decodificação por contagem de frequência de caracteres.

Assim, analisando o algoritmo de forma resumida, adotaremos os parâmetros listados em (1), (2) e (3) para codificar a mensagem em (4) e, em seguida, tomaremos outro parâmetro em (5) para decodificar a mesma em (6). É importante destacar que decodificar consiste em passar a mensagem codificada à sua forma original, ao contrário, o método não serviria.

5.3 Exemplos

Para ilustrar o algoritmo, apresentaremos exemplos didáticos nos quais é possível enxergar a validade do método utilizado, ou seja, vamos descumprir o primeiro passo de escolher primos *p* e *q* com mais de 100 algarismos, pois isso dificultaria uma compreensão mais imediata do fato.

Exemplo 1 – Utilize o algoritmo RSA para encriptar a palavra ALUNO.

(0) A palavra ALUNO, convertida em números, resulta em 1021302324;

(1) Escolhe-se $p = 3$ e $q = 11$;

(2) Calcula-se $n = 3 \cdot 11 = 33$ e $\phi(33) = (3-1) \cdot (11-1) = 2 \cdot 10 = 20$;

(3) Escolhe-se um valor para *e* relativamente primo com $\phi(33) = 20$. Escolheremos $e = 7$;

(4) Chamamos o par $(33,7)$ de *chave de codificação* do sistema RSA. Agora, a mensagem deve ser codificada em blocos *b* (consideremos que cada bloco tenha dois algarismos).

Denotaremos cada bloco codificado por $C(b)$, de modo que $C(b)$ é o resto da divisão de b^7 por *n*, ou seja, $C(b) \equiv b^7 \pmod{33}$. Cada bloco *b* será dado pelo número correspondente à conversão feita em (0). Daí, vem:

Tabela 3 – Encriptação da palavra ALUNO

Símbolo	Bloco numérico correspondente	b^7	$C(b) \equiv b^7 \pmod{33}$	Bloco numérico após encriptação
A	10	10^7	$C(10) \equiv 10^7 \pmod{33}$	10
L	21	21^7	$C(21) \equiv 21^7 \pmod{33}$	21
U	30	30^7	$C(30) \equiv 30^7 \pmod{33}$	24
N	23	23^7	$C(23) \equiv 23^7 \pmod{33}$	23
O	24	24^7	$C(24) \equiv 24^7 \pmod{33}$	18

Fonte: Elaborado pela autora

Logo, a palavra ALUNO encriptada transforma-se em 1021242318.

Para constatar a validade do método, vamos tentar desencriptar a palavra. Assim:
(5) Escolhe-se um inteiro positivo *d*, de modo que $d \cdot 7 \equiv 1 \pmod{20}$. Logo, por tentativa, tomemos $d = 3$;
(6) Chamamos o par $(33,3)$ de *chave de decodificação* do sistema RSA. Agora, a mensagem deve ser decodificada em blocos *a*. Denotaremos cada bloco decodificado por $D(a)$, de modo que $D(a)$ é o resto da divisão de a^3 por *n*, ou seja, $D(a) \equiv a^3 \pmod{33}$. Daí, vem:

Tabela 4 – Desencriptação da palavra ALUNO

Bloco numérico após encriptação	a^3	$D(a) \equiv a^3 \pmod{33}$	Bloco numérico após desencriptação
10	10^3	$D(10) \equiv 10^3 \pmod{33}$	10
21	21^3	$D(21) \equiv 21^3 \pmod{33}$	21
24	24^3	$D(24) \equiv 24^3 \pmod{33}$	30
23	23^3	$D(23) \equiv 23^3 \pmod{33}$	23

Bloco numérico após encriptação	a^3	$D(a) \equiv a^3 (\mathrm{mod}\, 33)$	Bloco numérico após desencriptação
18	18^3	$D(18) \equiv 18^3 (\mathrm{mod}\, 33)$	24

Fonte: Elaborado pela autora

Logo, ao desencriptar, voltamos a 1021302324 que corresponde à palavra ALUNO.

Exemplo 2 – Utilize o algoritmo RSA para encriptar a palavra PROFMAT.

(0) A palavra PROFMAT convertida em números, resulta em 25272415221029;
(1) Escolhe-se $p = 11$ e $q = 13$;
(2) Calcula-se $n = 11 \cdot 13 = 143$ e $\phi(143) = (11-1) \cdot (13-1) = 10 \cdot 12 = 120$;
(3) Escolhe-se um valor para *e* relativamente primo com $\phi(143) = 120$. Escolheremos $e = 7$;
(4) Chamamos o par $(143, 7)$ de *chave de codificação* do sistema RSA. Agora, a mensagem deve ser codificada em blocos *b* (consideremos que cada bloco tenha dois algarismos). Denotaremos cada bloco codificado por $C(b)$, de modo que $C(b)$ é o resto da divisão de b^7 por *n*, ou seja, $C(b) \equiv b^7 (\mathrm{mod}143)$. Cada bloco *b* será dado pelo número correspondente à conversão feita em (0). Daí, vem:

Tabela 5 – Encriptação da palavra PROFMAT

Símbolo	Bloco numérico correspondente	b^7	$C(b) \equiv b^7 (\mathrm{mod}143)$	Bloco numérico após encriptação
P	25	25^7	$C(25) \equiv 25^7 (\mathrm{mod}143)$	64
R	27	27^7	$C(27) \equiv 27^7 (\mathrm{mod}143)$	14
O	24	24^7	$C(24) \equiv 24^7 (\mathrm{mod}143)$	106
F	15	15^7	$C(15) \equiv 15^7 (\mathrm{mod}143)$	115
M	22	22^7	$C(22) \equiv 22^7 (\mathrm{mod}143)$	22
A	10	10^7	$C(10) \equiv 10^7 (\mathrm{mod}143)$	10
T	29	29^7	$C(29) \equiv 29^7 (\mathrm{mod}143)$	94

Fonte: Elaborado pela autora

Logo, a palavra PROFMAT encriptada transforma-se em 6414106115221094.

Para constatar a validade do método, vamos tentar desencriptar a palavra. Assim:
(5) Escolhe-se um inteiro positivo *d*, de modo que $d \cdot 7 \equiv 1 (\text{mod} 120)$. Logo, temos $d = 103$;
(6) Chamamos o par $(143, 103)$ de *chave de decodificação* do sistema RSA. Agora, a mensagem deve ser decodificada em blocos *a*. Denotaremos cada bloco decodificado por $D(a)$, de modo que $D(a)$ é o resto da divisão de a^{103} por *n*, ou seja, $D(a) \equiv a^{103} (\text{mod} 143)$. Daí, vem:

Tabela 6 – Desencriptação da palavra PROFMAT

Bloco numérico após encriptação	a^{103}	$D(a) \equiv a^{103} (\text{mod} 143)$	Bloco numérico após desencriptação
64	64^{103}	$D(64) \equiv 64^{103} (\text{mod} 143)$	25
14	14^{103}	$D(14) \equiv 14^{103} (\text{mod} 143)$	27
106	106^{103}	$D(106) \equiv 106^{103} (\text{mod} 143)$	24
115	115^{103}	$D(115) \equiv 115^{103} (\text{mod} 143)$	15
22	22^{103}	$D(22) \equiv 22^{103} (\text{mod} 143)$	22
10	10^{103}	$D(10) \equiv 10^{103} (\text{mod} 143)$	10
94	94^{103}	$D(94) \equiv 94^{103} (\text{mod} 143)$	29

Fonte: Elaborado pela autora

Logo, ao desencriptar, voltamos a 25272415221029, que corresponde à palavra PROFMAT.

Exemplo 3 – Utilize o algoritmo RSA para encriptar a frase CRIPTOGRAFIA É ARTE.

(0) A frase CRIPTOGRAFIA É ARTE, convertida em números, resulta em 1227182529241627101518109914991027291 4;
(1) Escolhe-se $p = 11$ e $q = 13$;
(2) Calcula-se $n = 11 \cdot 13 = 143$ e $\phi(143) = (11-1) \cdot (13-1) = 10 \cdot 12 = 120$;
(3) Escolhe-se um valor para *e* relativamente primo com $\phi(143) = 120$. Escolheremos $e = 7$;

(4) Chamamos o par $(143, 7)$ de *chave de codificação* do sistema RSA. Agora, a mensagem deve ser codificada em blocos *b* (consideremos que cada bloco tenha um, dois ou três algarismos de modo que cada bloco é menor que o número 143). Denotaremos cada bloco codificado por $C(b)$, de modo que $C(b)$ é o resto da divisão de b^7 por *n*, ou seja, $C(b) \equiv b^7 (\text{mod} 143)$. Cada bloco *b* será dado pelo número correspondente à conversão feita em (0). Então, da divisão em blocos, temos:

122 – 71 – 82 – 52 – 92 – 41 – 62 – 7 – 101 – 51 – 8 – 109 – 91 – 49 – 9 – 102 – 72 – 9 – 14

Assim sendo, segue a encriptação:

Tabela 7 – Encriptação da frase CRIPTOGRAFIA É ARTE

Bloco numérico correspondente	$C(b) \equiv b^7 (\text{mod} 143)$	Bloco numérico após encriptação
122	$C(122) \equiv 122^7 (\text{mod} 143)$	34
71	$C(71) \equiv 71^7 (\text{mod} 143)$	124
82	$C(82) \equiv 82^7 (\text{mod} 143)$	69
52	$C(52) \equiv 52^7 (\text{mod} 143)$	13
92	$C(92) \equiv 92^7 (\text{mod} 143)$	27
41	$C(41) \equiv 41^7 (\text{mod} 143)$	24
62	$C(62) \equiv 62^7 (\text{mod} 143)$	127
7	$C(7) \equiv 7^7 (\text{mod} 143)$	6
101	$C(101) \equiv 101^7 (\text{mod} 143)$	62
51	$C(51) \equiv 51^7 (\text{mod} 143)$	116
8	$C(8) \equiv 8^7 (\text{mod} 143)$	57
109	$C(109) \equiv 109^7 (\text{mod} 143)$	21
91	$C(91) \equiv 91^7 (\text{mod} 143)$	130
49	$C(49) \equiv 49^7 (\text{mod} 143)$	36
9	$C(9) \equiv 9^7 (\text{mod} 143)$	48

Bloco numérico correspondente	$C(b) \equiv b^7 (\text{mod}143)$	Bloco numérico após encriptação
102	$C(102) \equiv 102^7 (\text{mod}143)$	119
72	$C(72) \equiv 72^7 (\text{mod}143)$	19
9	$C(9) \equiv 9^7 (\text{mod}143)$	48
14	$C(14) \equiv 14^7 (\text{mod}143)$	53

Fonte: Elaborado pela autora

Logo, a mensagem CRIPTOGRAFIA É ARTE, encriptada, transforma-se em 341246913272412766211657211303648119194853.

Para constatar a validade do método, vamos tentar desencriptar a mensagem. Assim:
(5) Escolhe-se um inteiro positivo *d*, de modo que $d \cdot 7 \equiv 1(\text{mod}120)$. Logo, temos $d = 103$;
(6) Chamamos o par $(143,103)$ de *chave de decodificação* do sistema RSA. Agora, a mensagem deve ser decodificada em blocos *a*. Denotaremos cada bloco decodificado por $D(a)$, de modo que $D(a)$ é o resto da divisão de a^{103} por *n*, ou seja, $D(a) \equiv a^{103}(\text{mod}143)$. Portanto, temos:

Tabela 8 – Desencriptação da frase CRIPTOGRAFIA É ARTE

Bloco numérico após encriptação	$D(a) \equiv a^{103}(\text{mod}143)$	Bloco numérico após desencriptação
34	$D(34) \equiv 34^{103}(\text{mod}143)$	122
124	$D(124) \equiv 124^{103}(\text{mod}143)$	71
69	$D(69) \equiv 69^{103}(\text{mod}143)$	82
13	$D(13) \equiv 13^{103}(\text{mod}143)$	52
27	$D(27) \equiv 27^{103}(\text{mod}143)$	92
24	$D(24) \equiv 24^{103}(\text{mod}143)$	41
127	$D(127) \equiv 127^{103}(\text{mod}143)$	62
6	$D(6) \equiv 6^{103}(\text{mod}143)$	7

Bloco numérico após encriptação	$D(a) \equiv a^{103} (\text{mod} 143)$	Bloco numérico após desencriptação
62	$D(62) \equiv 62^{103} (\text{mod} 143)$	101
116	$D(116) \equiv 116^{103} (\text{mod} 143)$	51
57	$D(57) \equiv 57^{103} (\text{mod} 143)$	8
21	$D(21) \equiv 21^{103} (\text{mod} 143)$	109
130	$D(130) \equiv 130^{103} (\text{mod} 143)$	91
36	$D(36) \equiv 36^{103} (\text{mod} 143)$	49
48	$D(48) \equiv 48^{103} (\text{mod} 143)$	9
119	$D(119) \equiv 119^{103} (\text{mod} 143)$	102
19	$D(19) \equiv 19^{103} (\text{mod} 143)$	72
48	$D(48) \equiv 48^{103} (\text{mod} 143)$	9
53	$D(53) \equiv 53^{103} (\text{mod} 143)$	14

Fonte: Elaborado pela autora

Logo, ao desencriptar, voltamos a 12271825292416271015181099149910272914, que corresponde à frase CRIPTOGRAFIA É ARTE.

5.4 A fatoração de inteiros e a quebra de códigos

Pudemos perceber nos exemplos considerados no tópico anterior que o algoritmo nos possibilita encriptar uma palavra e obter a desencriptação correspondente. O trabalho, apesar de apresentar alguns cálculos mais extensos, é contornável com o auxílio das congruências. Por outro lado, nas transações com uso da *internet*, deve-se contar com o auxílio de um computador, pois aí necessita-se de um grau de segurança muito maior do que aquele apresentado nos exemplos.

Além disso, é importante lembrar que a codificação é feita considerando o número *n* que é produto de dois primos e a decodificação é feita utilizando exatamente os primos que se obtém ao fatorar *n*. Portanto, a decodificação baseia-se na determinação dos fatores de *n*. Assim, apresentaremos a seguir algoritmos para fatorar *n* e algumas aplicações.

5.4.1 Fatorando n

O primeiro método que utilizaremos para fatorar *n* consiste em tomar $n = pq$, com *p* e *q* primos e $\phi(n) = (p-1)\cdot(q-1)$. Daí, vem que: $\phi(n) = (p-1)\cdot(q-1) = pq - p - q + 1 = n - p - q + 1$. Desse modo, temos: (1): $p + q = n + 1 - \phi(n)$ e (2): $(p+q)^2 - 4pq = p^2 + 2pq + q^2 - 4pq = p^2 - 2pq + q^2 = (p-q)^2$. Substituindo (1) em (2): $(n+1-\phi(n))^2 - 4n = (p-q)^2 \Rightarrow p - q = \sqrt{(n+1-\phi(n))^2 - 4n}$. Logo, seguem as equações: $p + q = n + 1 - \phi(n)$ e $p - q = \sqrt{(n+1-\phi(n))^2 - 4n}$, e destas duas equações vem: $p = \dfrac{\sqrt{(n+1-\phi(n))^2 - 4n} + n + 1 - \phi(n)}{2}$ e $q = \dfrac{-\sqrt{(n+1-\phi(n))^2 - 4n} + n + 1 - \phi(n)}{2}$. Assim, temos a fatoração de *n*.

Para o segundo método, vamos supor inicialmente que o número *n* que se quer fatorar é ímpar, pois se for par, o número 2 é um de seus fatores. Assim, estamos interessados em determinar inteiros positivos tais que $n = x^2 - y^2 = (x+y)\cdot(x-y)$, ou seja, tais que $x + y$ e $x - y$ sejam fatores de *n*. Para determiná-los, vamos aplicar o algoritmo de Fermat e para isso devemos calcular a parte inteira da raiz quadrada de *n*. Usaremos a simbologia $\lfloor \sqrt{n} \rfloor$. Por exemplo: $\lfloor \sqrt{17} \rfloor = 4$ e $[\pi] = 3$. Quando *n* é um quadrado perfeito, $n = r^2$. Assim, *r* é um fator de *n* e, na notação acima, $x = r$ e $y = 0$. Segue o algoritmo:

Algoritmo de Fermat:

(1) Calculamos $x = \lfloor \sqrt{n} \rfloor$. Se $n = x^2$, então *x* é um fator de *n* e podemos parar;

(2) Se *x* não é um fator de *n*, incrementamos *x* de uma unidade e calculamos $y = \sqrt{x^2 - n}$;

(3) Repetir a etapa (2) até encontrar um valor inteiro para *y* ou até que *x* seja igual a $\dfrac{n+1}{2}$.

Quando encontramos um *y* inteiro, o número *n* tem fatores $x + y$ e $x - y$. Quando $x = \dfrac{n+1}{2}$, o número *n* é primo.

Demonstração. Para demonstrar o funcionamento do algoritmo, é necessário considerar separadamente o que acontece quando *n* é composto e quando *n* é primo. No primeiro caso, precisamos mostrar que existe um inteiro $x > \lfloor \sqrt{n} \rfloor$ tal que $\sqrt{x^2 - n}$ é um inteiro menor que $\frac{n+1}{2}$. Isto significa que se *n* é composto, então o algoritmo para antes de chegar a $\frac{n+1}{2}$. Se *n* é primo, então é necessário verificar que o único valor de *x* possível é $\frac{n+1}{2}$.

Suponhamos que *n* pode ser fatorado na forma $n = ab$, onde $a \leq b$. Assim, queremos obter inteiros positivos *x* e *y* tais que $n = x^2 - y^2$, ou seja, $n = ab = (x-y)\cdot(x+y) = x^2 - y^2$. Como $x - y \leq x + y$, isto sugere que tomemos $a = x - y$ e $b = x + y$. Resolvendo as equações na forma de sistema, temos: $x = \frac{b+a}{2}$ e $y = \frac{b-a}{2}$, onde $\left(\frac{b+a}{2}\right)^2 - \left(\frac{b-a}{2}\right)^2 = ab = n$. Note que *x* e *y* têm que ser inteiros, mas estão escritos na forma de fração. Porém, *n* é ímpar, por hipótese. Logo, *a* e *b*, fatores de *n*, também são ímpares. Daí, temos que as frações consideradas para *x* e *y* são números inteiros e, por isso, precisamos supor que a entrada do algoritmo é sempre um número ímpar.

Se *n* é primo, só podemos ter $a = 1$ e $b = n$ ou vice versa. Assim, $x = \frac{n+1}{2}$ é o único valor para *x* quando *n* é primo. No caso de *n* ser composto, devemos considerar dois casos: i) Se $a = b$, o algoritmo obtém a resposta desejada já na etapa (1) do algoritmo; ii) Se *n* é composto e não é um quadrado perfeito, o algoritmo vai parar se forem satisfeitas as desigualdades $\lfloor \sqrt{n} \rfloor \leq \frac{b+a}{2} < \frac{n+1}{2}$.

Lembremos que a variável *x* é inicializada com o valor $\lfloor \sqrt{n} \rfloor$ e que vai sendo incrementada de uma unidade a cada laço. Assim, as desigualdades anteriores nos garantem que se *n* for composto, chegamos a $\frac{b+a}{2}$ antes de chegar a $\frac{n+1}{2}$. Quando $x = \frac{b+a}{2}$, $y^2 = \left(\frac{b+a}{2}\right)^2 - n = \left(\frac{b-a}{2}\right)^2$. Atingindo este laço, o algoritmo para, obtendo *a* e *b* como fatores. Portanto, se *n* é composto, o algoritmo sempre para antes de chegar a $x = \frac{n+1}{2}$, tendo conseguido determinar fatores de *n*.

Uma vez provada a validade dos métodos apresentados, podemos destacar que ambos são bem interessantes: o primeiro admite que podemos determinar *p* e *q*, a partir de *n* e de $\phi(n)$; já o segundo – o algoritmo de Fermat – recebe um número *n* e devolve sua fatoração em primos *p* e *q* ou a mensagem que diz que *n* é um número primo.

Mesmo que, aparentemente, seja complicado considerar a utilização nos métodos apresentados de números com uma quantidade grande de algarismos, é importante destacar que a escolha de primos não deve levar em conta apenas o tamanho dos números, até porque, nos dias de hoje, podemos contar com diversas ferramentas computacionais e programas que podem executar os métodos de fatoração demonstrados, de modo a favorecer a submissão de números considerados inviáveis para o trabalho básico com papel, lápis e uma calculadora convencional que porventura estejamos interessados em executar. Sendo assim:

> lembre-se que a segurança do RSA depende da dificuldade de fatorar a chave pública n, que é igual ao produto de dois primos. A primeira impressão é que basta escolher os primos grandes, para garantir que n é difícil de fatorar. Mas isto *não* é verdade. Por exemplo, se escolhermos os primos grandes, mas próximos, então n é facilmente fatorável pelo algoritmo de Fermat. (COUTINHO, 2013, p. 43)

No entanto, independentemente de sua funcionalidade, os métodos de fatoração esbarram no problema do custo computacional. Apesar de poderem ser implementados, os métodos descritos apresentam um custo computacional bastante elevado para certas fatorações. Deste modo, não vale a pena submeter determinados números compostos para ser fatorados com o uso de algum dos métodos mesmo que essa submissão conte com a fatoração feita por meio dos computadores mais potentes. Se isso fosse feito, os programas iriam funcionar por anos e ainda assim a fatoração de alguns números seria tida por impossível.

Portanto, uma vez que entendemos que em determinadas condições a fatoração em primos (mesmo com uma quantidade grande de algarismos), apesar de trabalhosa, é possível, apresentaremos alguns exemplos em que esses métodos podem ser utilizados sem maiores problemas.

5.4.2 Exemplos

Exemplo 1 (Disponível em Coutinho 2013, p. 191) – Sabendo-se que $n = 3552377$ é igual ao produto de dois números primos e que $\phi(n) = 3548580$; fatore *n*.

Resolução. Utilizaremos o primeiro método apresentado, através do qual conseguimos determinar os fatores *p* e *q* de *n*, a partir de *n* e de $\phi(n)$. Assim, basta-nos substituir os valores em $p = \frac{\sqrt{(n+1-\phi(n))^2 - 4n} + n + 1 - \phi(n)}{2}$ e $q = \frac{-\sqrt{(n+1-\phi(n))^2 - 4n} + n + 1 - \phi(n)}{2}$. Assim, com os valores dados, teremos: $\sqrt{(n+1-\phi(n))^2 - 4n} = 464$. Daí, $p = \frac{464 + 3798}{2} = \frac{4262}{2} = 2131$ e $q = \frac{-464 + 3798}{2} = \frac{3334}{2} = 1167$. Portanto, temos que: $n = 2131 \times 1167$.

Exemplo 2 (Disponível em Cavalcante [s.d], p. 6 – Adaptado) Ao utilizar um algoritmo RSA, adotou-se $n = 9981401593$ (extenso); fatore *n*.

Resolução. Vamos utilizar o algoritmo de Fermat. Inicialmente, determinaremos a parte inteira da raiz quadrada de *n*. Portanto, $x = \lfloor \sqrt{9981401593} \rfloor = 99906$. Seguimos incrementando *x* em uma unidade para poder determinar o valor inteiro de *y* ou até que *x* seja igual a $\frac{n+1}{2}$ que, no caso do exemplo, é igual a 49900700797. Nesse caso, incrementando *x* em uma unidade, já conseguimos determinar um *y* inteiro. Para $x = 99907$, temos $y = 84$. Logo, os fatores de *n* são $x + y = 99907 + 84 = 99991$ e $x - y = 99907 - 84 = 99823$. Então, $n = 99991 \times 99823$.

Exemplo 3 (Disponível em Coutinho 2013, p. 191) – A mensagem 6355 – 5075 foi codificada pelo método RSA usando a senha $n = 7597$ e $e = 4947$. Além disso, sabe-se que $\phi(n) = 7420$. Decodifique a mensagem.

Resolução. Para decodificar a mensagem, devemos determinar a chave de decodificação, ou seja, o par (n, d). De acordo com algoritmo apresentado no tópico 5.2, a desencriptação é feita seguindo os passos (5) e (6) apresentados. Teremos:

(5) Escolhe-se um inteiro positivo *d*, de modo que $d \cdot e \equiv 1 (\mathrm{mod}\, \phi(n))$, ou seja, $d \cdot 4947 \equiv 1 (\mathrm{mod}\, 7420)$. Logo, por tentativa, temos $d = 3$;

(6) Chamamos o par $(7597, 3)$ de *chave de decodificação* do sistema RSA. Agora, a mensagem deve ser decodificada em blocos *a*. Denotaremos cada bloco decodificado por $D(a)$, de modo que $D(a)$ é o resto da divisão de a^3 por *n*, ou seja, $D(a) \equiv a^3 (\text{mod } 7597)$.

Portanto, temos:

$$D(6355) \equiv 6355^3 (\text{mod } 7597)$$
$$6355^2 \equiv 373 (\text{mod } 7597)$$
$$6355^3 \equiv 373 \cdot 6355 (\text{mod } 7597) \equiv 151 (\text{mod } 7597)$$

$$D(5075) \equiv 5075^3 (\text{mod } 7597)$$
$$5075^2 \equiv 1795 (\text{mod } 7597)$$
$$5075^3 \equiv 1795 \cdot 5075 (\text{mod } 7597) \equiv 822 (\text{mod } 7597)$$

Ao desencriptar a mensagem 6355 – 5075, obtemos 151 – 822. Utilizando a correspondência entre letras e números de dois algarismos sugeridos no início do capítulo, temos 15 – 18 – 22, que corresponde à palavra FIM.

5.5 Funcionamento do Algoritmo RSA

Para que o algoritmo base da criptografia RSA funcione, é preciso, conforme já mencionado, que ao decodificar uma mensagem, obtenhamos a mensagem original. Se isso sempre ocorrer, o método é válido, do contrário, o método de criptografia não fará sentido. A esse respeito, Coutinho (2013, p. 183) destaca:

> [...] é claro que se *b* é um bloco da mensagem original, então esperamos que *D(C(b))=b*. Em outras palavras, decodificando um bloco da mensagem codificada, esperamos encontrar o bloco correspondente da mensagem original. Sem isto, não temos nenhum código útil. Entretanto, não é óbvio que isto é sempre verdade [...]

Deste modo, precisamos garantir que o algoritmo sempre funciona, e é o que faremos na propriedade a seguir.

Propriedade 5.5.1 Denotando cada bloco codificado por $C(b)$ e cada bloco decodificado por $D(a)$, aplicando o processo de decodificação a um bloco codificado, obtemos o bloco correspondente da mensagem original, ou seja, $D(C(b)) = b$.

Demonstração. De acordo com o algoritmo, temos que os blocos codificado e decodificado correspondem, respectivamente, a $C(b) \equiv b^e \pmod{n}$ e $D(a) \equiv a^d \pmod{n}$. Assim, para decodificar uma mensagem codificada, fazemos: $D(C(b)) \equiv C(b)^d \equiv b^{ed} \pmod{n}$.Como *d* é o inverso de *e* módulo $\phi(n)$, temos que $d \cdot e \equiv 1 (\mathrm{mod}\, \phi(n))$, ou seja, $d \cdot e = 1 + k\phi(n)$, onde *k* é um inteiro. Daí, segue que $D(C(b)) \equiv b^{1+k\phi(n)} \pmod{n} \equiv \left(b^{\phi(n)}\right)^k \cdot b \pmod{n}$. Como $n = pq$, temos $\phi(n) = (p-1)\cdot(q-1)$, o que implica que $D(C(b)) \equiv \left(b^{(p-1)}\right)^{(q-1)\cdot k} \cdot b \pmod{n}$. Se *p* não divide *b*, pelo pequeno teorema de Fermat, $b^{p-1} \equiv 1 \pmod{p}$, então obtemos $b^{ed} \equiv b \pmod{p}$, ou seja, $D(C(b)) \equiv b \pmod{p}$. Se, ao contrário, *p* divide *b*, então $b \equiv 0 \pmod{p}$, ou seja, $\left(b^{(p-1)}\right)^{(q-1)\cdot k} \cdot b \equiv 0 \pmod{p}$. Analogamente, é possível mostrar que $D(C(b)) \equiv b \pmod{q}$ e como *p* e *q* são primos, temos que $mdc(p,q) = 1$, sendo que *pq* divide $b^{ed} - b$. Como $n = pq$, concluímos que $D(C(b)) \equiv b \pmod{n}$.

De tal forma, de posse da validade da proposição, garantia para o funcionamento do algoritmo, apresentaremos no tópico a seguir os motivos que fazem o método RSA ser tão seguro.

5.6 Segurança

É importante lembrar que o RSA é um modelo de criptografia assimétrica, ou seja, o funcionamento do método consiste na determinação de duas chaves. Funciona assim: codifica-se uma mensagem utilizando uma chave A, denominada chave pública. Todas as pessoas podem ter acesso a essa chave, pois ela é, de fato, pública. Uma vez que a mensagem é enviada, mesmo o detentor da chave pública não terá acesso à mesma. A decodificação da mensagem fica a cargo de quem possuir a chave B ou chave privada. Cada chave pública gera uma chave privada específica. Deste modo, a chave privada é de domínio bem restrito e isso dificulta a decifração por terceiros, que teriam que quebrar o código para lerem a mensagem enviada.

Matematicamente, sendo *p* e *q* os parâmetros do sistema usado para o algoritmo RSA, com $n = pq$, a chave pública necessita de um inteiro *e* relativamente primo com $\phi(n) = (p-1)\cdot(q-1)$, pois o par (n,e) representa a chave pública ou chave de codificação. Já a chave privada depende de um inteiro *d*, tal que $de \equiv 1(\mathrm{mod}\,\phi(n))$. O par (n,d) corresponde à chave privada ou chave de decodificação. Portanto, o RSA só será seguro se for difícil calcular *d* a partir de *n* e de *e*. O cálculo de *d* necessita de *e* e de $\phi(n)$. O valor $\phi(n)$, por outro lado, só pode ser calculado se soubermos fatorar *n* para obter *p* e *q*.

É válido destacar que a escolha dos primos *p* e *q* que geram *n* é de grande importância. Se os valores forem pequenos, o sistema pode ser quebrado facilmente; se forem grandes, com uma diferença $|p-q|$ pequena, como o exemplo 2 de 5.4.2, o algoritmo de Fermat poderá capturá-los. A esse respeito, Coutinho (2013, p. 187) relata:

> isto não é puro papo furado. Em 1995 dois estudantes de uma universidade americana quebraram uma versão do RSA em uso público. O feito só foi possível porque a escolha dos primos usada neste sistema era inteiramente inadequada. Por outro lado, o RSA está em uso há anos e, se os primos forem bem escolhidos, o sistema tem-se mostrado muito seguro. Portanto, uma receita para escolher primos bons é essencial para a "caixa de ferramentas" de qualquer um que deseje programar o RSA.

Portanto, uma vez feita uma boa escolha de primos, somos levados a crer que a segurança do RSA é garantida pelo alto custo computacional, que gera a impossibilidade da fatoração de um número composto *n* considerado. Até hoje, não temos métodos de baixo custo que agilizem tal fatoração.

6 CONSIDERAÇÕES FINAIS

Não foi por acaso que a evolução da criptografia trouxe a segurança que as transações da *internet* necessitavam. De acordo com o exposto, pudemos observar que a criptografia assimétrica (da qual a RSA é o principal exemplo) é mais segura que a simétrica. Por possuir duas chaves, a dificuldade de decodificar uma mensagem codificada com a RSA é bem maior do que naquelas mensagens secretas que utilizavam chave única. Apesar de ser mais lenta, a criptografia assimétrica não necessita de uma comunicação inicial entre emissor e receptor e isso, entre outros fatores, a torna mais segura.

À medida que passamos a compreender as noções empregadas no algoritmo RSA e nos apropriamos das mesmas, somos levados a perceber que quanto maior for o tamanho da chave empregada na codificação e decodificação, mais segura estará a informação transmitida. Dessa forma, temos que a escolha do tamanho da chave de codificação deverá levar em conta o grau de importância e o tamanho da informação que se queira proteger. Se a informação for de alta relevância, como um arquivo que contém as senhas bancárias dos correntistas ou o número de seu cartão de crédito, precisa-se de uma criptografia mais forte, devido à importância das informações. Mas se for de baixa relevância, como um arquivo pessoal, pode-se utilizar chaves menores para criptografar, por que será mais rápido.

Além disso, vale destacar que a garantia de segurança do algoritmo tem por base chaves originadas por um par de primos bem escolhidos, ou seja, os números primos apresentam-se como elementos de importância inquestionável para a segurança dos computadores e das informações transmitidas. Sautoy (2007, p. 332-333) completa:

> a história dos primos se estende muito além do mundo matemático. Os avanços tecnológicos mudaram nossa maneira de fazer matemática. O computador, nascido em Bletchley Park, nos deu a capacidade de enxergar números que antigamente ficavam confinados no universo inobservável. [...] O papel central dos primos na segurança dos computadores jogou esses números ao centro das atenções. Atualmente, os primos afetam as vidas de todos nós, pois protegem os segredos eletrônicos mundiais do olhar impertinente dos *hackers* da internet.

Assim, os átomos da aritmética continuam a fazer o seu papel na segurança do método que sustenta toda nossa "vida virtual". Sem os primos e o método criado a partir de sua utilização, o comércio e as transações realizadas pela *internet*, que proporcionam uma facilidade de interação que torna nossa vida tão cômoda, seriam impossíveis. Sautoy (2007, p.

333) nos lembra que embora os segredos dos primos permaneçam ocultos, a importância prática que eles adquiriram não pode deixar de ser considerada. Além disso, vale destacar que

> nosso consolo é que, mesmo que os primos jamais revelem seus segredos, estão nos levando a uma odisseia intelectual extraordinária. A importância que adquiriram se estende muito além de seu papel fundamental como os átomos da aritmética. Como descobrimos, os primos abriram portas entre áreas até então desconexas da matemática. (SAUTOY, 2007, p. 334-335)

Assim, a Teoria dos Números que tratava, como o próprio nome sugere, de teoria, viu-se inserida em um mundo prático que anseia por mais descobertas. Mesmo assim:

> apesar dos grandes esforços feitos pelas maiores mentes matemáticas na tentativa de explicar a modulação e a transformação dessa música mística, os primos ainda são um enigma sem resposta. Aguardamos a pessoa cujo nome viverá para sempre como o matemático que despertou a canção dos primos. (SAUTOY, 2007, p. 335)

Portanto, até que o enigma dos primos seja descoberto, nossas transações comerciais estão seguras. O problema não resolvido da fatoração em primos abriu as portas para um mundo onde segurança é a palavra de ordem. A confiança com a qual depositamos nossas informações está baseada justamente na impossibilidade de solução do problema, ou seja, enquanto o problema da fatoração não puder ser resolvido, o algoritmo e os dados depositados na grande rede estão protegidos. Ao contrário, um novo algoritmo precisa ser implementado com urgência.

BIBIOGRAFIA

CAMPELLO, Antônio Carlos e Leal, Isabel. **Teoria Aritmética dos Números e Criptografia RSA**. Disponível em: <http://www.ime.unicamp.br/~ftorres/ENSINO/MONOGRAFIAS/antonio_RSA.pdf>. Acesso em: 29 de março de 2015.

CAVALCANTE, André L. B. **Teoria dos Números e Criptografia**. UPIS Faculdades Integradas – Faculdade de Tecnologia. Departamento Sistemas de Informação. Disponível em: <file:///C:/Users/Lana/Downloads/teoria%20dos%20n-meros%20e%20criptografia.pdf>. Acesso em: 29 de março de 2015.

COUTINHO, S. C. **Números Inteiros e Criptografia RSA**. Rio de Janeiro: IMPA, 2013.

EVES, Howard. **Introdução à história da Matemática**. Campinas, SP: Editora da UNICAMP, 2004.

GIL, Antônio Carlos. **Métodos e Técnicas de Pesquisa Social**. 6ª Edição. São Paulo: Editora Atlas S.A., 2008. (p. 49-59)

HEFEZ, A. **Elementos de Aritmética**. 2ª Edição. Textos Universitários: SBM, 2005.

HEFEZ, A. **Aritmética**. 1ª Edição. SBM, 2013 (Coleção PROFMAT).

JOGO da imitação, O. Título original: *The Imitation Game*. Direção: Morten Tyldum. Reino Unido / EUA: Diamond Filmes, 2014. 114 min.

KLÉO, J. S. C. **Teoria dos Números: semestre III**. Fortaleza: UAB/IFCE, 2010.

KRISCHER, Thais Cristine. **Um estudo da Máquina Enigma**. 2013. 98 f. Trabalho de Conclusão de Curso (Graduação em Ciências da Computação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2013. Disponível em: < https://www.lume.ufrgs.br/bitstream/handle/10183/66106/000870987.pdf?sequence=1>. Acesso em: 19 de março de 2015.

LOVÁSZ, L.; PELIKÁN J.; VESZTERGOMBI, K. **Matemática Discreta**. 2ª Edição. Rio de Janeiro: SBM, 2013.

MUNIZ NETO, Antonio Caminha. **Tópicos de Matemática Elementar: Teoria dos Números.** 2ª Edição. Rio de Janeiro: SBM, 2013.

OLIVEIRA, Krerley M. O. e FERNANDÉZ, Adán J. C. **Iniciação à Matemática: um curso com problemas e soluções**. 2ª Edição. Rio de Janeiro: SBM, 2010.

RIBENBOIM, Paulo. **Números Primos: Velhos Mistérios e Novos Recordes**. 1ª Edição. Rio de Janeiro: IMPA, 2012.

SANTOS, José Plínio de O. **Introdução à Teoria dos Números**. 3ª Edição. Rio de Janeiro: IMPA, 2012.

SAUTOY, Marcus Du. **A música dos números primos: a história de um problema não resolvido na matemática**. Rio de Janeiro: ZAHAR, 2007.

SILVA, Elen V. Pereira. **Introdução à Criptografia RSA**. Universidade Estadual Paulista – UNESP. Departamento de Matemática. Ilha Solteira, São Paulo, 2006. Disponível em: <http://www.impa.br/opencms/pt/eventos/downloads/jornadas_2006/trabalhos/jornadas_elen_pereira.pdf >. Acesso em: 22 de março de 2015.