



DETERMINAÇÃO DA CONSTANTE DE HUBBLE E IDADE DO UNIVERSO ATRAVÉS DA OBSERVAÇÃO DE LENTES GRAVITACIONAIS

RELATÓRIO FINAL DE PROJETO DE INICIAÇÃO CIENTÍFICA (PIBIC/INPE/CNPq)

Alex Oliveira Rocha (Mackenzie, Bolsista PIBIC/CNPq)
E-mail: rocha.alex.oliveira@gmail.com

Luiz Claudio Lima Botti (CRAAM, Centro de Radioastronomia e Astrofísica
Mackenzie, Escola de Engenharia, Universidade Presbiteriana
Mackenzie/INPE/DAS(DIDAS)/CEA/MCTI e ROPK, Orientador)
E-mail: luizquas@yahoo.com.br

JULHO DE 2020

DETERMINAÇÃO DA CONSTANTE DE HUBBLE E IDADE DO UNIVERSO ATRAVÉS DA OBSERVAÇÃO DE LENTES GRAVITACIONAIS

RESUMO

As lentes gravitacionais foram previstas por Einstein na Teoria da Relatividade Geral em 1915, que diz que o campo gravitacional de um objeto de massa elevada encurva a geometria local do Universo, de forma que a luz ao passar perto do objeto também se encurva.

Este trabalho tem como objetivo a determinação da constante de Hubble e a Idade do Universo através da observação de Lentes Gravitacionais.

Este projeto iniciado em março de 2019, tem como objetivo utilizar-se da variabilidade destes objetos utilizando dados obtidos de curvas de luz para determinar atrasos de sinais entre componentes de diversos quasares utilizando algoritmos e técnicas matemáticas.

Os sistemas que foram estudados neste projeto foram os quasares-lentes: HE0435-1223, HS 2009+1914, RX J1131-1231, SDSS J1001+5027 e SDSS J1206+4332.

Utilizando dois métodos diferentes para se obter o *time delay* das componentes dos quasares, sendo um deles o método de Interpolação Harmônica de Fourier e o outro utilizando Redes Neurais Recorrentes, foi possível estimar a constante de Hubble sabendo-se que a mesma é proporcional a velocidade de dispersão da galáxia-lente e inversamente proporcional ao atraso de sinais de suas componentes. Dessa forma, sabendo-se o atraso entre os sinais A e B de duas componentes ou mais dos quasares, pode-se calcular a constante de Hubble. Como o inverso da mesma possui dimensão de tempo, pode-se desta maneira estimar a idade do universo.

DETERMINATION OF THE HUBBLE CONSTANT AND AGE OF THE UNIVERSE THROUGH THE OBSERVATION OF GRAVITATION LENSES

ABSTRACT

Gravitational lenses were predicted by Einstein in the Theory of General Relativity in 1915, which says that the gravitational field of an object of high mass bends the local geometry of the Universe, so that the light when passing close to the object also bends.

This work aims to determine the Hubble constant and the Age of the Universe through the observation of Gravitational Lenses.

This project started in March 2019, aims to use the variability of these objects using data obtained from light curves to determine signal delays between components of different quasars using mathematical algorithms and techniques.

The systems that were studied in this project were the quasar-lenses: HE0435-1223, HS 2009 + 1914, RX J1131-1231, SDSS J1001 + 5027 and SDSS J1206 + 4332.

Using two different methods to obtain the time delay of the quasar components, one of them being the Fourier Harmonic Interpolation method and the other using Recurrent Neural Networks, it was possible to estimate the Hubble constant knowing that it is proportional to the speed dispersion of the galaxy-lens and inversely proportional to the delay of signals from its components. Thus, knowing the delay between signals A and B of two or more components of the quasars, one can calculate the Hubble constant. As the inverse of it has a dimension of time, it is possible to estimate the age of the universe.

AGRADECIMENTOS

Em primeiro lugar, ao meu orientador Luiz Claudio Lima Botti pela amizade, compreensão, confiança, orientação, parceria e sempre estar a disposição para me ajudar contribuindo para o desenvolvimento do projeto e ao meu amadurecimento acadêmico.

Aos colegas e pesquisadores do Centro de Radioastronomia do Mackenzie (CRAAM) que contribuíram transmitindo seus conhecimentos oriundos de suas áreas de atuação.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela oportunidade e fomento durante o período de vigência do contrato junto à instituição.

À Universidade Presbiteriana Mackenzie (UPM), pelo espaço e a infraestrutura cedida para o desenvolver do projeto.

Ao projeto COSMOGRAIL, pelo fornecimento dos dados que foram utilizados nesta pesquisa publicamente.

À minha família, pelo amor, incentivo e apoio incondicional durante o desenvolvimento do projeto.

À minha namorada – Letícia, por sempre me apoiar e ajudar durante todos os momentos.

Finalmente, agradeço ao Instituto Nacional de Pesquisas Espaciais (INPE/PIBIC/MCTIC) pela oportunidade cedida.

ÍNDICE DE ILUSTRAÇÕES

Lista de Figuras

<i>Figura 1: Efeito Lente Gravitacional</i>	10
<i>Figura 2: Anel de Einstein</i>	11
<i>Figura 3: Cruz de Einstein</i>	11
<i>Figura 4: HE0435-1223</i>	13
<i>Figura 5: HS 2209+1914</i>	14
<i>Figura 6: RX J1131-1231</i>	15
<i>Figura 7: SDSS J1001+5027 e SDSS J1206+4332</i>	16
<i>Figura 8: Espectro eletromagnético</i>	17
<i>Figura 9: Efeito Doppler Sonoro</i>	18
<i>Figura 10: Imagem do Quasar Duplo 0957+561 separados a 6"</i>	20
<i>Figura 11: Diagrama simplificado de um Neurônio Biológico</i>	22
<i>Figura 12: Um neurônio simples</i>	24
<i>Figura 13: Estrutura típica de uma rede neural multicamada</i>	25

Lista de Tabelas

<i>Tabela 1: Comparação entre as componentes do RX J1131-1231</i>	50
<i>Tabela 2: Resultados</i>	60

Lista de Gráficos

<i>Gráfico 1: Dados do HE0435-1223</i>	37
<i>Gráfico 22: Componente A do HE0435-1223</i>	38
<i>Gráfico 3: Componente B do HE0435-1223</i>	38
<i>Gráfico 4: Componente C do HE0435-1223</i>	39
<i>Gráfico 5: Componente D do HE0435-1223</i>	39
<i>Gráfico 6: Componentes A e D do candidato a lente gravitacional HE0435-1223</i>	40
<i>Gráfico 7: Determinação do time delay entre os picos de máximo e de mínimo do HE0435-1223</i>	41
<i>Gráfico 8: Pontos criados pela RNR do objeto HE0435-1223</i>	42
<i>Gráfico 9: HS 2209+1914</i>	43
<i>Gráfico 10: Componente A do HS 2209+1914</i>	43
<i>Gráfico 11: Componente B do HS 2209+1914</i>	44
<i>Gráfico 12: Análise das Curvas de Luz do HS 2209+1914</i>	45
<i>Gráfico 13: Análise das Curvas de Luz do HS 2209+1914</i>	45
<i>Gráfico 14: Pontos criados pela RNR do objeto HS2209+1914</i>	46
<i>Gráfico 15: RX J1131-1231</i>	47
<i>Gráfico 16: Componente A do RX J1131-1231</i>	47
<i>Gráfico 17: Componente B do RX J1131-1231</i>	48

Gráfico 18: Componente C do RX J1131-1231	48
Gráfico 19: Componente D do RX J1131-1231	49
Gráfico 20: Análise das Curvas de Luz do RX J1131-1231	49
Gráfico 21: Análise das Curvas de Luz do RX J1131-1231	50
Gráfico 22: Pontos criados pela RNR do objeto RX J1131-1231	51
Gráfico 23: SDSS J1001+5027	52
Gráfico 24: Componente A do SDSS J1001+5027	52
Gráfico 25: Componente B do SDSS J1001+5027	53
Gráfico 26: Análise das Curvas de Luz do SDSS J1001+5027	54
Gráfico 27: Análise das Curvas de Luz do SDSS J1001+5027	54
Gráfico 28: Pontos criados pela RNR do objeto SDSS J1001+5027	55
Gráfico 29: SDSS J1206+4332	56
Gráfico 30: Componente A do SDSS J1206+4332	56
Gráfico 31: Componente B do SDSS J1206+4332	57
Gráfico 32: Análise das Curvas de Luz do SDSS J1206+4332	57
Gráfico 33: Análise das Curvas de Luz do SDSS J1206+4332	58
Gráfico 34: Pontos criados pela RNR do objeto SDSS J1206+4332	59

Lista de Equações

Equação 1: Redshift para velocidades pequenas	19
Equação 2: Lei de Hubble	19
Equação 3	24
Equação 4: Função Sigmoide.....	26
Equação 5: Gerador de valor de saída de um nó da camada oculta	27
Equação 6: Erro	27
Equação 7: Gradiente de Erro na forma diferencial	27
Equação 8: Gradiente de Erro	28
Equação 9: Gradiente de Erro para cada nó j.....	28
Equação 10: Atualização do peso entre camada de entrada e camada oculta.....	28
Equação 11: Atualização do peso entre camada oculta e camada de saída	28
Equação 12: Atualização do peso entre camada de entrada e camada oculta melhorada ..	29
Equação 13: Atualização do peso entre camada oculta e camada de saída melhorada	29
Equação 14: Função da Tangente Hiperbólica	30
Equação 15: Constante de Hubble	32
Equação 16: Idade do Universo	32
Equação 17	34
Equação 18: Estimativa do Redshift para desvios pequenos ($z < 1$)	41
Equação 19: Redshift para grandes desvios ($z > 1$)	41

SUMÁRIO

1.0 INTRODUÇÃO	8
2.0 OBJETIVOS	12
3.0 LENTES GRAVITACIONAIS ESTUDADAS	13
3.1 HE0435-1223	13
3.2 HS 2209+1914	14
3.3 RX J1131-1231	14
3.4 SDSS J1001+5027	15
3.5 SDSS J1206+4332	15
4.0 REVISÃO DA LITERATURA	17
4.1 O ESPECTRO ELETROMAGNÉTICO	17
4.2 HISTÓRICO DAS LENTES GRAVITACIONAIS	20
4.3 REDES NEURAIS	22
4.3.1 NEURÔNIOS BIOLÓGICOS	22
4.3.2 NEURÔNIOS ARTIFICIAIS	23
4.3.3 REDES NEURAIS EM MULTICAMADAS	25
4.3.4 AJUSTE DE PESOS EM REDES NEURAIS MULTICAMADAS	26
4.3.4.1 MÉTODO DA RETROPROPAGAÇÃO	26
4.3.5 REDES NEURAIS RECORRENTES	30
5.0 METODOLOGIA	32
5.1 MÉTODO DE ANÁLISE DAS CURVAS DE LUZ	34
6.0 CURVAS DE LUZ ANALISADAS	36
6.1 HE0435-1223	37
7.0 RESULTADOS	60
8.0 CONCLUSÕES	61
9.0 REFERÊNCIAS	62

1.0 INTRODUÇÃO

No século XVII, o conhecimento sobre os princípios fundamentais que regem a natureza era muito pequeno, de modo que a maioria das pessoas desta época acreditava que os fenômenos eram obras de um ser supremo que controlava todos os acontecimentos do Universo. Foi então em 1666 que Sir Isaac Newton (1642-1727) chegou enfim a uma teoria que esclarecia esta questão e muitas outras, a Teoria da Gravidade (DINIZ, 2009).

A Teoria da Gravidade diz que dois objetos quaisquer se atraem por meio de uma força que depende das massas desses objetos e da distância que há entre eles. Dados dois corpos com massa m e M , a uma distância d entre si, a força de atração desses corpos é mútua, proporcional ao produto das massas e inversamente proporcional ao quadrado dessa distância d . Porém, a teoria de Newton assume que alterações na força gravitacional são transmitidas imediatamente, contradizendo o fato de que há uma velocidade limite, a da luz no vácuo, e nada pode ter maior velocidade do que essa. Essa teoria também não explica ondas gravitacionais, que foram comprovadas em 11 de fevereiro de 2016 por cientistas do Observatório de Ondas Gravitacionais por Interferômetro Laser (ABBOTT, B. P. et al., 2016) e buracos negros. De acordo com a teoria Newtoniana, se o Universo é euclidiano, de densidade uniforme e infinito, a força gravitacional em um certo ponto é uma teoria que é incompatível com um Universo que possui estas propriedades (DINIZ, 2009).

Em 1783 foi mencionado o fato de que a gravidade poderia interferir num raio de luz. Este fenômeno foi discutido por John Michell (geólogo britânico e astrônomo) numa carta endereçada a Henry Cavendish (físico e químico britânico). Michell em 1784 estudou hipoteticamente o caso de um objeto celeste compacto e denso que poderia impedir que a luz escapasse da sua superfície. Ele pensou em um objeto invisível, mas que poderia ser identificado pelo movimento das estrelas próximas a ele. Laplace em 1795, com um intervalo de dez anos após a carta de John Michell, sugeriu a ideia de objetos com alta gravidade capturando a luz. A Teoria de Newton já contemplava a ideia de uma estrela escura, com altíssima gravidade, que levou ao conceito atualmente aceito de buraco negro.

Em 1784 Cavendish calculou a deflexão da luz por um corpo tendo por base a Teoria Corpuscular da Luz e a Teoria Newtoniana. Em 1804 John Soldner calculou o desvio de posição angular de uma estrela (Soldner, 1804). O desvio calculado foi de aproximadamente $0,84''$. Valores bem próximos foram calculados por Cavendish e por Einstein (1911) para o caso da interação da luz de uma estrela com o Sol.

As contradições da teoria de Newton levaram os cientistas e pensadores a procurar uma nova teoria, e então, em 1915 Albert Einstein (1879-1955) publicou sua Teoria da Relatividade Geral. A teoria de Einstein prediz que o campo gravitacional de um objeto de massa elevada encurva a geometria local do Universo, de forma que a luz ao passar perto do objeto também se encurva. Ao invés de descrever o efeito de gravitação como uma força, Einstein introduziu o conceito de espaço-tempo curvo. Einstein disse então que este encurvamento poderia ser verificado analisando os raios de luz provenientes de objetos próximos ao Sol, onde se constata que a posição de uma estrela, por exemplo, sofre deslocamento ao passar pelo Sol, o que de fato foi confirmado em 1919 quando ocorreu um eclipse Solar em Sobral no Ceará (EDDINGTON, 1920).

Após o desenvolvimento da Relatividade Geral em 1915, Einstein (1936) refez seus cálculos e chegou ao valor aproximado de 1,7” para o desvio de um raio de luz de uma estrela pelo campo gravitacional do Sol. A.S. Eddington foi um dos poucos a entender a Teoria da Relatividade no início do século passado e em 1919 observou o eclipse solar em Sobral (Ceará) e comprovou que Einstein estava certo, calculando as posições das estrelas no momento do eclipse e comparando com as posições à noite (Eddington, 1920a e 1920b).

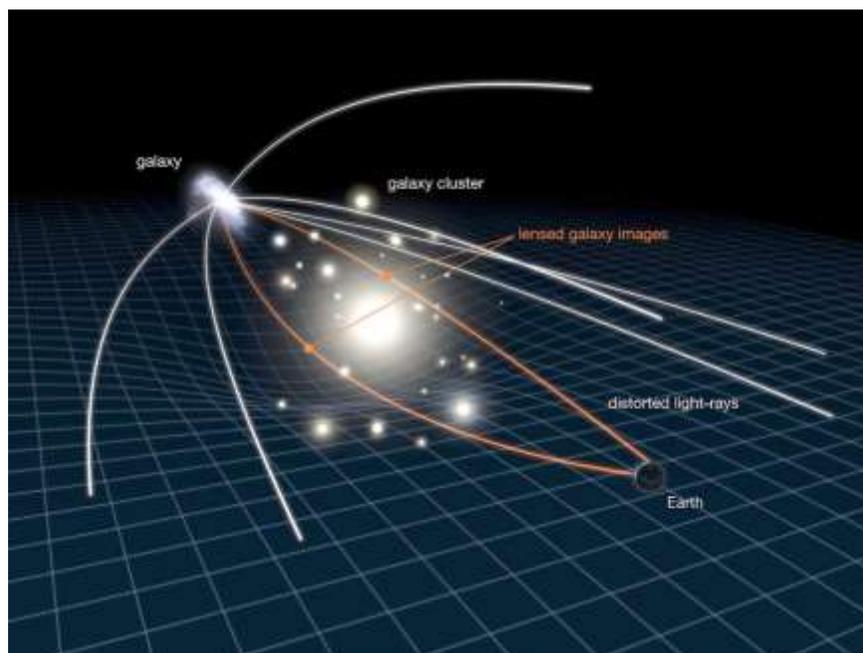
No eclipse as estrelas apresentaram-se mais afastadas uma das outras. O termo “Lente” foi proposto primeiramente por Oliver Lodge em 1919, para o desvio gravitacional da luz (Lodge, 1919). Em 1912, Einstein já tinha concebido a ideia de que um corpo massivo poderia defletir a luz proveniente de um objeto mais afastado, mas não a publicou. Somente em 1936 o fez a pedido de Rudi W. Mandl (Einstein, 1936). Em 1924, Orest Danilovich Chwolson concebeu a ideia de que duas estrelas perfeitamente alinhadas vistas por um observador também alinhado a elas, veria a imagem de um anel entorno da estrela que estaria no primeiro plano. Esses anéis foram chamados “Anéis de Chwolson”, mais tarde denominados “Anéis de Einstein”. Fritz Zwicky percebeu em 1930 que estes efeitos poderiam também ocorrer em grandes distâncias e neste caso galáxias poderiam ser compactas o suficiente para encurvar a luz de objetos situados a distâncias ainda maiores. Zwicky sugeriu que aglomerados de galáxias poderiam ser lentes gravitacionais. Somente em 1979 este efeito foi observado (Zwicky, 1937).

Na década de sessenta a ideia de lentes gravitacionais ressurgiu e 3 pesquisadores independentemente, Yu.G.Klimov, S.Liebes e S.Refsdal analisaram o efeito de galáxias funcionarem como lentes para outras galáxias. Estudaram o efeito de lente de estrelas em estrelas, de estrelas de aglomerados globulares em estrelas da Via Láctea, e também considerou-se a luz de uma supernova em uma galáxia distante (Klimov, 1963). A luz dessa supernova pode seguir trajetórias diferentes para o observador e haveria uma diferença de tempo entre os dois trajetos até o atingir. Daí pensou-se que a diferença de tempo entre estes dois trajetos poderia ser capaz de determinar a constante de Hubble e a massa da galáxia (Liebes (1964) e Refsdal (1964 a e b). Cooke e Kantowski (1975) elaboraram uma expressão geral para o tempo de atraso para várias imagens.

Lente gravitacional é um fenômeno no qual trajetórias de luz vindas de uma fonte são encurvadas por um potencial gravitacional localizado entre a fonte e o observador. Embora as imagens apareçam distintas da fonte, a Teoria da Relatividade Geral mostra uma clara relação entre as propriedades das imagens e a distribuição de massa das lentes. Este fato tem possibilitado o uso de lentes gravitacionais em Cosmologia, principalmente para calcular a distribuição de massa do Universo e estimar distâncias cosmológicas. O efeito de lente gravitacional tem sido estudado desde 1937 (Tikhov, 1937). Nesta época Tichov estudou o efeito causado pela deflexão gravitacional da luz de uma estrela “S” no campo gravitacional de outra estrela “B” (“S” está atrás da estrela “B” mas próxima à linha de base). Tikhov (1937) mostrou que um aumento considerável da luminosidade aparente de “S” poderia ser possível.

A Figura 1 mostra um campo gravitacional causado por um aglomerado de galáxias (Galaxy Cluster) que encurva os raios de luz oriundos de uma outra (X) que se situa do aglomerado de galáxias em relação ao observador, causando um efeito de múltiplas imagens da galáxia X.

Figura 1: Efeito Lente Gravitacional



Fonte: NASA/ESA

A primeira observação de uma lente gravitacional ocorreu apenas em 1979, quando foram vistas duas imagens do mesmo quasar (WALSH, CARSWELL E WEYMANN, 1979).

O efeito das Lentes Gravitacionais pode ocorrer em todas as escalas - de aglomerados de galáxias até pequenas estrelas e planetas, até mesmo ocorre quando raios de luz passam perto de nossos próprios corpos, porém este efeito é tão pequeno que é difícil de se medir. A quantidade e a forma das imagens do quasar dependem das posições deste e do objeto que funciona como lente e da posição do observador. Caso todos eles estejam perfeitamente alinhados, a imagem resultante será um anel, conhecido como “Anel de Einstein” (Figura 2).

Outra forma especial do efeito de Lente Gravitacional é a Cruz de Einstein, que corresponde a 4 imagens do mesmo objeto que aparecem em torno de uma galáxia em primeiro plano, graças a uma poderosa lente gravitacional. Na realidade, a estrutura é formada por cinco pontos, o ponto central e mais outros quatro pontos que correspondem a

uma imagem de uma galáxia que “quebra” a luz de um único objeto distante (no caso da Figura 3, um quasar) situado atrás em quatro pontos separados, na forma de cruz.

Figura 2: Anel de Einstein



Fonte: NASA / ESA

Figura 3: Cruz de Einstein



Fonte: NASA / ESA

2.0 OBJETIVOS

O presente projeto de pesquisa busca verificar qual é o comportamento temporal (curvas de luz) de diversos quasares candidatos a sistemas de lentes gravitacionais em diversas frequências do espectro eletromagnético. Mais especificamente, analisar qual seria a variabilidade de cada imagem-lente de um sistema, se existe correlação entre possíveis eventos e qual a importância de uma lente gravitacional para a Cosmologia.

Foram estudadas diversas curvas de luz associadas às imagens-lentes nesses sistemas de lentes gravitacionais e determinado o atraso entre os sinais de duas ou mais componentes utilizando-se algoritmos computacionais.

Do conhecimento deste atraso e da velocidade de dispersão da galáxia-lente foi calculada a constante de Hubble. Conhecendo-se a mesma foi possível estimar a idade do Universo.

Além disto, foi estudado a fundo redes neurais, que será revisto em teoria no próximo capítulo para aplicação aos dados obtidos das lentes gravitacionais, realizando o melhor ajuste possível às curvas para obtenção dos pontos faltantes, determinando assim o *time-delay* de cada fonte, conseqüentemente a constante de Hubble e a idade do Universo.

Uma observação importante é que foram realizados dois métodos diferentes para a obtenção do *time-delay* de cada fonte, o primeiro foi a partir de uma série de Fourier para ajustar os dados e o outro utilizando redes neurais.

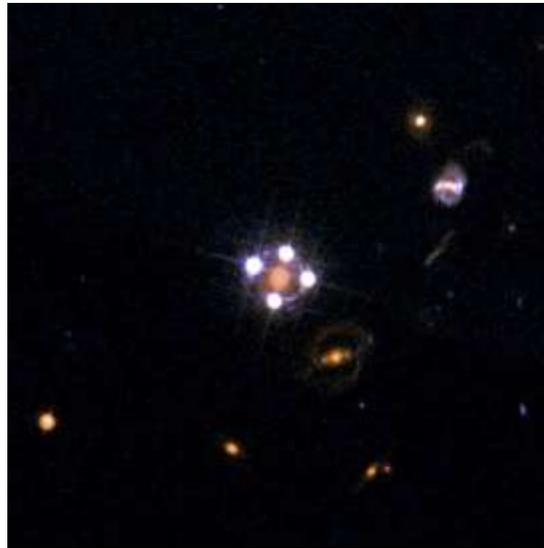
3.0 LENTES GRAVITACIONAIS ESTUDADAS

3.1 HE0435-1223

HE0435-1223 é um quasar-lente quádruplo, descoberto em 2002 por um estudo conjunto de Wisotzki et al., 2002. Este QSO possui quatro componentes em um arranjo em forma de cruz em torno de uma galáxia brilhante em seu centro (Figura 4), e a separação máxima entre as imagens é de 2,6 arcseg, permitindo uma decomposição confiável do sistema. Seu perfil espacial é o de uma galáxia elíptica com uma escala de 12 Kpc. O QSO mostra variabilidade, com magnitudes totais da banda *G* de 17,89 e 17,71 para duas épocas separadas por aproximadamente 2 meses. Entretanto, os fluxos relativos das componentes não mudam, indicando que as variações são intrínsecas.

Este sistema quasar-lente está entre as cinco lentes gravitacionais mais estudadas atualmente. A galáxia em primeiro plano cria quatro imagens quase uniformemente distribuídas do quasar distante em torno dele (Figura 4).

Figura 4: HE0435-1223



Fonte: ESA/Hubble, NASA, Suyu et al.

Como parte do projeto COSMOGRAIL, foi apresentado uma nova medida obtida neste trabalho para constante de Hubble (H_0) e outros parâmetros cosmológicos com base na análise conjunta de três sistemas e seus respectivos *time delays*. Com os mesmos dados que serão utilizados nesta pesquisa, foi observado 13 anos de medidas deste objeto, e foi obtido uma constante de Hubble de $H_0 = 71,9 \pm 1,8$ km/s/Mpc (V. Bonvin, arXiv:1607.01790 [astro-ph.CO]).

O *redshift* (conceito que será abordado no capítulo 2.1) deste quasar-lente foi medido como $z = 1,689$ por Wisotzki et al., 2002.

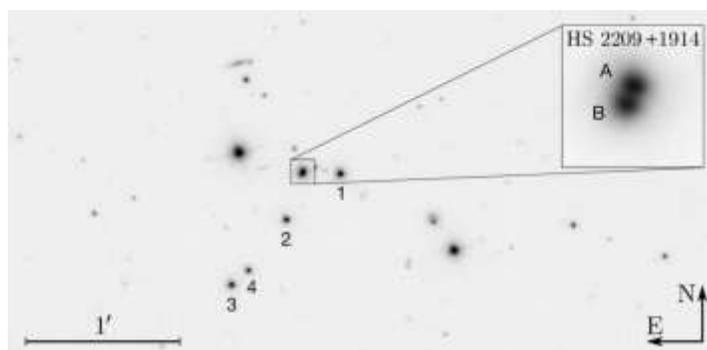
3.2 HS 2209+1914

HS 2209+1914 é um sistema de quasar-lente duplo que foi observado durante o período de 2004 a 2011 na banda *R*.

Com a colaboração do projeto COSMOGRAIL, na literatura já se apresentam valores para o *time delay* deste sistema, que foi de $\Delta t = 20,0 \pm 5$ dias do fluxo da imagem de B para A utilizando técnicas como método da dispersão, spline cúbico, mínimos quadrados, entre outros Eulaers et al., 2013.

Localizada na Figura 5, é uma lente gravitacional que cria duas imagens do quasar distante.

Figura 5: HS 2209+1914



Fonte: Mercator Telescope/Cosmograil

3.3 RX J1131-1231

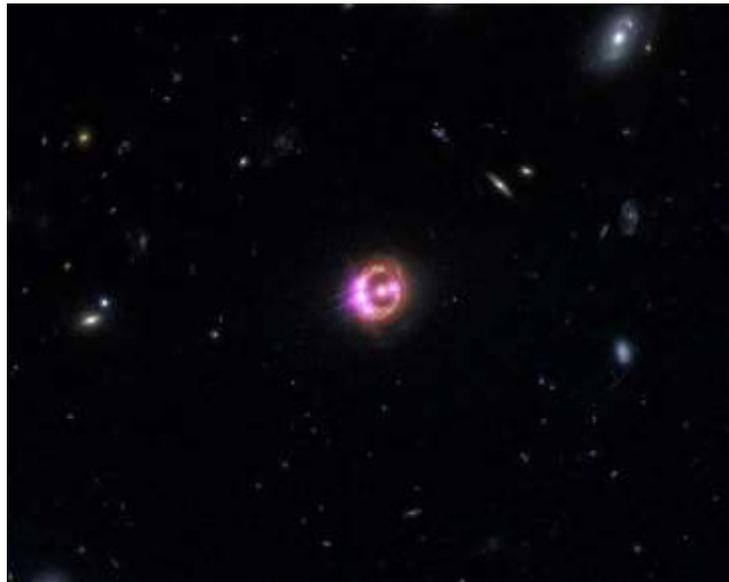
É um distante buraco negro supermaciço contendo um quasar localizado a cerca de 6 bilhões de anos luz da Terra na constelação Taça. Pode-se observar este objeto no centro da Figura 6, onde quatro imagens foram criadas em torno do quasar distante.

Este sistema está envolvido com a descoberta em 2014 que raios X provenientes de uma região dentro do disco de acreção de um buraco negro localizado cerca de três vezes o raio do horizonte de eventos. Isso implicava que o buraco negro deveria girar incrivelmente rápido para permitir que o disco sobreviva em um raio tão pequeno. Com isto, pode-se realizar a primeira medida direta da rotação de um buraco negro.

Esta determinação foi feita por uma equipe liderada por Rubens Reis da Universidade de Michigan usando o observatório de raios-x Chandra da NASA e os telescópios da estação espacial européia.

No entanto, as medidas não poderiam ter ocorrido sem um raro alinhamento do quasar e uma galáxia elíptica gigante (que é parte de um aglomerado de outras galáxias alinhadas com o quasar) que fica entre a Terra e o RX J1131-1231 (REDD, 2014). Essa formação forneceu uma lente gravitacional quádrupla que ampliou a luz proveniente do quasar.

Figura 6: RX J1131-1231



Fonte: NASA

3.4 SDSS J1001+5027

SDSS J1001+5027 é um sistema de lente gravitacional duplo, com um *redshift* $z = 1,838$ (Kumar,2013), há dados deste QSO por mais de 6 anos, entre março de 2005 e julho de 2011, usando o telescópio MERCATOR. Foi detectado uma forte variabilidade intrínseca neste quasar, com uma amplitude na ordem de magnitude de 0,2.

Na literatura, foi observado um *time delay* da imagem A para a imagem B de $119,3 \pm 3,3$ dias (Kumar, 2013)

3.5 SDSS J1206+4332

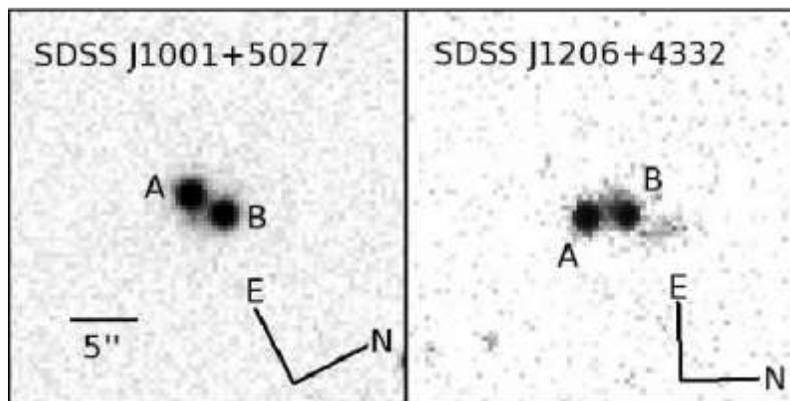
SDSS J1206+4332 também é um sistema de lente gravitacional duplo, monitorado assim como o anterior na banda *R* usando quatro telescópios: Mercator, Maidanak,

Himalayan Chandra e Euler Telescopes, juntos observando este sistema durante o período de 2004 e 2011 (Eulaers, E., 2013).

Na literatura, foi encontrado um *time delay* da imagem A para a imagem B de $111,3 \pm 3$ dias.

Localizada na Figura 7, são lentes gravitacionais que criam duas imagens do quasar distante.

Figura 7: SDSS J1001+5027 e SDSS J1206+4332



Fonte: Mercator Telescope/Cosmograil

Todos os dados das curvas de luz destes quasares foram obtidos através do projeto COSMOGRAIL.

O projeto COSMOGRAIL, da sigla de “*The COSmological MONitoring of GRAvItational Lenses XII*” consiste de um grupo de pesquisa da Universidade de Cornell que buscam monitorar candidatos à lentes gravitacionais de todo universo visível.

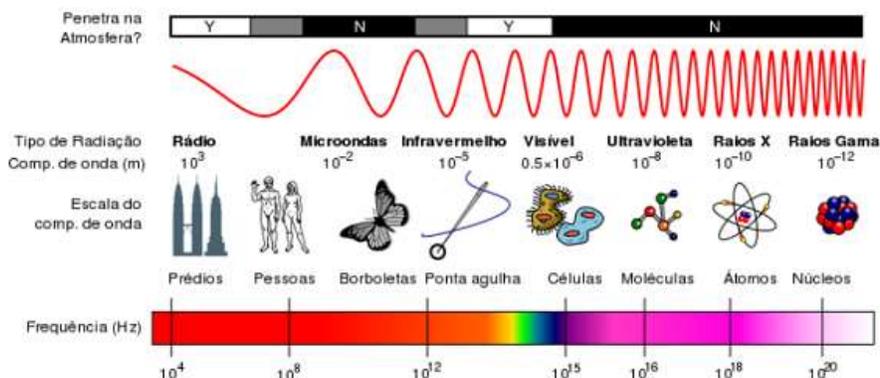
4.0 REVISÃO DA LITERATURA

4.1 O ESPECTRO ELETROMAGNÉTICO

Uma das formas de propagação da luz ocorre através de movimento ondulatório e ela não necessita de meios para se propagar, já que é uma onda de caráter eletromagnético. Uma partícula carregada em movimento cria um campo elétrico que oscila e se propaga no espaço, e este por sua vez gera um campo magnético que o acompanha (FRIAÇA et al. 2000).

O olho humano vê a luz que corresponde a apenas uma pequena parte do espectro eletromagnético. O espectro eletromagnético inclui as radiações rádio, infravermelho, ultravioleta, raios-X e raios gama, sendo que essas formas de radiação variam em seu comprimento de onda. A atmosfera terrestre bloqueia essas radiações, e o que passa por ela são apenas radiações na região do visível, pouco do infravermelho e parte de rádio, como pode ser observado na Figura 8.

Figura 8: Espectro eletromagnético



Fonte: Villate, Jaime E. Faculdade de Engenharia – Universidade do Porto

Como se pode observar, a luz visível corresponde a apenas uma fração de todo o espectro eletromagnético. À esquerda do espectro visível, encontra-se os tipos de energia de frequência mais baixa, que incluem radiação infravermelha, que são ondas de calor emitidas por corpos térmicos, e sua aplicação vai além de diversas outras áreas das ciências exatas, um bom exemplo é a comunicação entre o controle remoto e o aparelho de televisão dos modelos mais antigos.

A direita do espectro visível, pode-se observar os raios ultravioletas (UV), raios X e raios gama. Esses tipos de ondas são extremamente prejudiciais para os organismos vivos, uma vez que suas frequências são muito altas, conseqüentemente, uma grande quantidade de energia.

Na natureza pode-se encontrar mais de um tipo de espectro, sólidos ou líquidos aquecidos ou gases a alta pressão emitem luz em todos os comprimentos de onda, diz-se

então que emitem um espectro contínuo. A região onde o corpo emite luz predominante depende da temperatura do corpo. Quanto mais quente, menor o comprimento de onda da luz, conseqüentemente maior a frequência emitida.

Os espectros são características intrínsecas dos objetos, e a partir desses espectros é possível estudar a composição química das estrelas e o *redshift* de galáxias.

A onda de rádio, como a luz visível, é uma forma de radiação eletromagnética que se propaga na velocidade da luz. Ao contrário das ondas mecânicas que aparecem na superfície de um líquido ou de ondas sonoras, a radiação eletromagnética não necessita de um meio para se propagar, isto é, pode-se propagar no vácuo, o que permite observar objetos muito distantes, como estrelas e galáxias (DINIZ, 2009).

Em 1842, Christian Doppler (1803-1853) demonstrou que ao movimentar uma fonte de onda sonora o comprimento de onda percebido pelo observador em repouso era afetado. Ao se aproximar de uma onda sonora o comprimento de onda diminui, resultando em um som agudo, do contrário o comprimento de onda aumenta e o som é mais grave. Este efeito ficou conhecido como efeito Doppler. Uma situação pode ser imaginada, uma ambulância com sua sirene ligada desloca-se em uma rua afastando-se de um observador e aproximando-se de um outro observador, como pode-se observar na Figura 9.

Figura 9: Efeito Doppler Sonoro

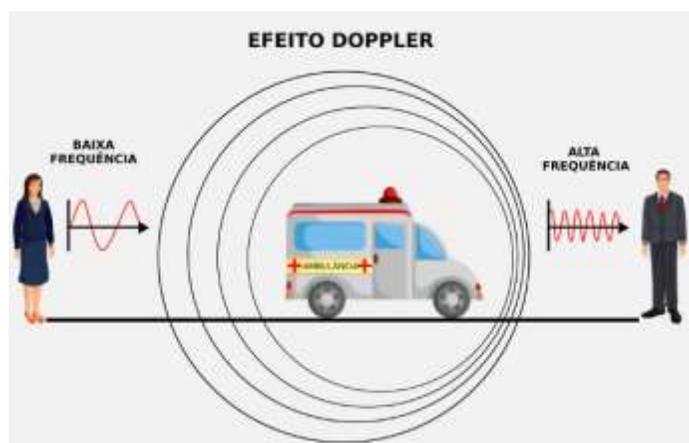


Ilustração: Rafael Helerbrock

O efeito Doppler também pode ser observado nas ondas eletromagnéticas, como a luz. Assim como no caso do som, a velocidade da luz não depende de seu observador, somente do meio no qual ela se propaga. Portanto, quando há aproximação entre a fonte de ondas eletromagnéticas e um observador, este perceberá um aumento nas frequências observadas e uma diminuição do comprimento de onda, e quando a fonte se afasta do observador, este perceberá uma diminuição nas frequências observadas, conseqüentemente um aumento no comprimento de onda.

Portanto, quando há aproximação entre a fonte e o observador, a frequência observada aumenta e fica mais próximo do “azul”. Os astrônomos chamam este fenômeno de *blue-shift*. Já quando a fonte se afasta do observador, a frequência diminui e fica mais próximo do “vermelho”. Este último é um dos fenômenos mais importantes para a Astronomia Extragaláctica, que é o *redshift*.

Em 1929, Edwin Powell Hubble (1889-1953) estudando o *redshift* das galáxias encontrou que suas velocidades de recessão eram proporcionais à distância da Terra, assim, os astrônomos puderam concluir que quanto mais distante o objeto, maior o deslocamento para o vermelho (FRIAÇA et al. 2000).

A expressão do *redshift* (z) para um objeto que se desloque de seu comportamento de onda (λ_0) para outro (λ) é dada pela equação (1):

$$z = \frac{\lambda_0 - \lambda}{\lambda} = \frac{v}{c}$$

Equação 1: Redshift para velocidades pequenas

Onde v é a velocidade de recessão e c é a velocidade da luz (FRIAÇA et al. 2000).

O *redshift* é então o resultado da expansão do universo, e a partir dele foi formulada a lei de Hubble, que é a medida de como se dá esta expansão, conforme a equação (2).

$$v = H_0 d$$

Equação 2: Lei de Hubble

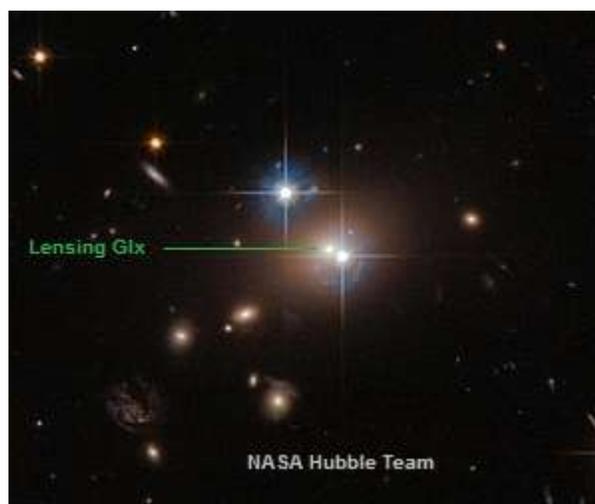
Sendo d a distância das galáxias e H_0 a constante de Hubble (FRIAÇA et al. 2000).

Deste modo, os quasares, que são os objetos que possuem os maiores *redshifts* são os objetos mais distantes do Universo e ao mesmo tempo são os que se formaram pouco tempo após o Big Bang.

4.2 HISTÓRICO DAS LENTES GRAVITACIONAIS

A primeira lente gravitacional foi descoberta por Walsh, Carswell e Weymann (1979). Analisando o espectro de cada imagem, pôde-se concluir que o quasar duplo *0957+561* era na verdade uma imagem dupla de um mesmo objeto, com o *redshift* $z = 1,41$. Após um ano da descoberta deste quasar, a galáxia lente foi identificada, uma gigante elíptica com um $z = 0,39$, situada ao norte da componente *B* deste quasar (COMBES et al. 2001) (Figura 10).

Figura 10: Imagem do Quasar Duplo *0957+561* separados a 6"



Fonte: NASA / ESA – Hubble

Desde a descoberta do *0957+561* a busca por casos similares se intensificou, e a segunda lente gravitacional foi descoberta por Weymann, 1980 apenas um ano após a primeira descoberta.

A lente gravitacional descoberta por Weymann foi a *PG 1115+080*, que era constituída de três imagens, A, B e C, e a componente A era muito mais brilhante que as outras duas, a componente B apresentava cerca de 10% do fluxo da A, e a C aproximadamente 18%. A componente C ficava mais distante de A que B, e a separação entre A e B é de aproximadamente $2,3''$, e por ter essa separação relativamente grande, esse sistema foi muito difícil de se observar. O *redshift* do quasar é cerca de 1,7 (SCHNEIDER et al, 1992).

O ganho na resolução espacial alcançada desde os anos oitenta permitiu a compreensão detalhada do sistema do quasar *PG 1115+080*, e apesar de ter sido descoberto como um sistema triplo, observações mais detalhadas revelaram que era na verdade um sistema quádruplo (YOUNG et al. 1981), confirmando então que a componente A era na verdade constituída de duas imagens, A1 e A2, tendo um espaço entre elas de aproximadamente 0,5 arcos-segundos. Ainda segundo Young et al. (1981), quando um

quasar está muito próximo a um centro de uma galáxia, deve-se resultar em cinco imagens do mesmo objeto, e uma geometria usual deste efeito é uma “cruz torta”, com quatro imagens brilhantes e uma fraca, e então a *PG 1115+080* teria, na realidade, cinco imagens, porém a quinta por ser muito fraca, até hoje não foi detectada.

Como o intervalo de dados observacionais disponíveis do *PG 1115+080* é muito amplo, ele é tomado como um exemplo para mostrar como as observações ajudam na determinação de parâmetros cosmológicos, como a constante de Hubble (H_0). O tempo de atraso encontrado na componente C em relação a A1 + A2 foi de 9,4 dias, de 23,7 dias para C em relação a componente B e A1+A2 em relação a B atrasa 14,3 dias. Foi adotado $\pm 3,4$ dias como a melhor estimativa para todos os atrasos (SCHECHETER et al. 1997).

O efeito de lente gravitacional que ocasiona múltiplas imagens é apenas um dos tipos dos efeitos conhecidos. Há outros tipos de efeitos, como o arco gravitacional, cruz de Einstein, anel de Einstein e lentes fracas, porém todos possuem o mesmo princípio.

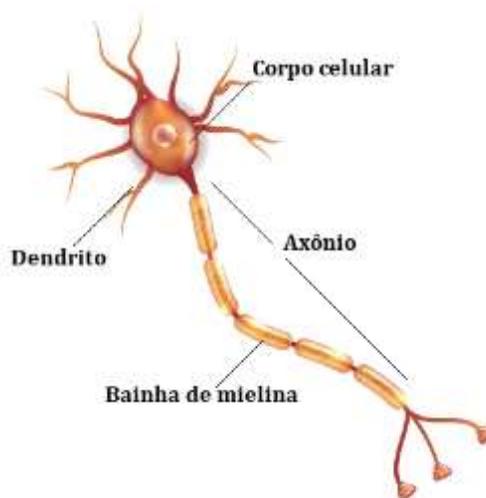
4.3 REDES NEURAIS

4.3.1 NEURÔNIOS BIOLÓGICOS

O cérebro humano possui mais de dez bilhões de neurônios, cada qual conectado, em média, a milhares de outros neurônios. Estas conexões são conhecidas por sinapses e o cérebro humano possui cerca de 60 trilhões destas conexões (COPPIN, 2015).

Neurônios são, na verdade, elementos de processamento muito simples. Cada neurônio é composto por uma soma, que é o corpo do neurônio, um axônio e vários dendritos. Pode-se observar um diagrama simplificado de um neurônio biológico na Figura 11.

Figura 11: Diagrama simplificado de um Neurônio Biológico



Fonte: Ilustração por Vanessa Sardinha dos Santos

O neurônio recebe entradas de outros neurônios, junto ao seus dendritos e quando este sinal de entrada excede um limiar o neurônio é *ativado* – na verdade, ocorre uma reação química que resulta em um pulso elétrico, conhecido como potencial de ação, que será enviado ao axônio (a saída do neurônio), em direção às sinapses que conectam um neurônio aos dendritos de outros neurônios.

Embora cada neurônio individualmente seja uma estrutura muito simples, quando integrado aos demais formam uma rede de extrema complexidade e esta é capaz de processar informação em uma extraordinária velocidade. Em termos de complexidade, o cérebro humano excede, de longe, qualquer dispositivo criado pelo homem, ou até mesmo qualquer outro objeto ou estrutura que ocorra naturalmente no universo, pelo o que é conhecido atualmente.

O cérebro humano usa uma forma de atribuição de crédito para reforçar conexões entre neurônios que levam a soluções corretas para seus problemas e enfraquecem conexões que levam a respostas incorretas. A intensidade de uma conexão, ou sinapse, determina qual será sua influência nos neurônios desta conexão e, assim, se uma conexão for enfraquecida, seu papel será atenuado nas próximas “computações” (COPPIN, 2015).

4.3.2 NEURÔNIOS ARTIFICIAIS

Uma rede neural constituída de neurônios artificiais é um sistema de processamento de informações que contém uma alta performance analisando características em comuns entre os dados. Este tipo de rede foi desenvolvido para generalizações de modelos matemáticos da cognição humana ou sua rede neural biológica, baseando as seguintes hipóteses:

- O processamento de informação ocorre através de um elemento muito simples chamado de neurônio.
- Sinais são transmitidos entre neurônios através de conexões.
- Cada conexão possui um peso associado no qual em um neurônio biológico multiplica o sinal transmitido.
- Cada neurônio possui uma função de ativação no qual realiza uma operação entre o sinal de entrada para obter o sinal de saída.

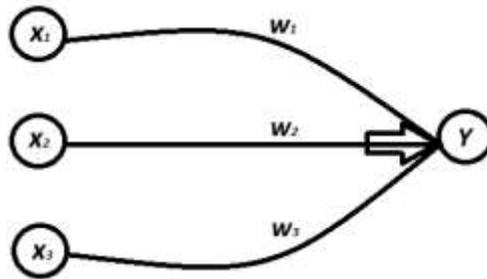
Uma rede neural artificial é caracterizada por (1) seu padrão de conexão entre os neurônios (chamada de arquitetura), (2) seu método de determinar os pesos entre as conexões (chamado de *treinamento*, ou *aprendizado*) e (3) sua função de ativação.

Redes neurais artificiais são moldadas com base no cérebro humano e são compostas por diversos neurônios artificiais, porém, como deve-se esperar, o número de neurônios artificiais nestas redes são todas significativamente menores do que no cérebro humano, conseqüentemente, haverá menos conexões entre eles.

Cada neurônio (ou nó) em uma rede neural (ou neuronal, como alguns autores o chamam) recebe uma série de entradas. Uma função chamada de função de ativação é aplicada a esses valores de entrada, o que resulta no nível de ativação do neurônio, que é o valor de saída dele. Há uma série de funções que podem ser utilizadas nos neurônios, como a função degrau, função sigmoide ou até mesmo uma função linear.

Um exemplo, considere um neurônio Y , ilustrado na Figura 12, que recebe sinais de entrada de outros três neurônios, X_1 , X_2 e X_3 . Os sinais de saída destes neurônios são x_1 , x_2 e x_3 respectivamente. O peso destas conexões entre os neurônios X_1 , X_2 e X_3 com o neurônio Y são w_1 , w_2 e w_3 respectivamente.

Figura 12: Um neurônio simples



Fonte: Acervo do Autor.

Uma função das mais comumente usadas é a função degrau ou a função limiar linear. Ao usar esta função, as entradas para o neurônio são somadas (tendo cada uma sido multiplicada por um peso) e esta soma é comparada com um limiar t . Se a soma for maior do que o limiar, o neurônio será ativado e terá um nível de ativação “+ 1”. Caso contrário, o neurônio ficará inativo e terá nível zero de ativação.

Então, o comportamento do neurônio pode ser expresso da através da Equação 3.

Equação 3

$$Y = \sum_{i=1}^n w_i x_i$$

Logo, pode-se observar que o comportamento do neurônio dependerá da soma ponderada entre os n sinais de entrada das conexões multiplicadas pelos seus pesos correspondentes. Utilizando o exemplo, considera-se que cada entrada associada ao neurônio Y tenha os seguintes pesos: $w_1 = 0.8$, $w_2 = 0.4$ e $w_3 = 0.6$ e os neurônios de entradas possuem $x_1 = 0.6$, $x_2 = 0.7$ e $x_3 = 0.9$. Então, a somatória entre os pesos e os sinais de entrada vale: $Y = (0.8 * 0.6) + (0.4 * 0.7) + (0.6 * 0.9) = 1.3$.

Suponha também que o nível de ativação do neurônio $f(x)$ é definido como:

$$f(x) = \begin{cases} +1, & Y > t \\ 0, & Y \leq t \end{cases}$$

Então, para o neurônio ativar com +1, deve-se ter um limiar t menor ou igual a 1.3, caso contrário, ele terá zero como nível de ativação e não será ativado.

Um neurônio que usa a função de ativação linear utilizará simplesmente a soma ponderada de suas entradas como nível de ativação. Também há a função sigmoide, que converte entradas na faixa de $-\infty$ a $+\infty$ em um nível de ativação que varia de 0 a +1.

Deste modo, quando há diversos destes neurônios conectados, produz uma rede neural que não possui processamento central ou mecanismo de controle, pois toda rede é

envolvida em cada fragmento de computação que ocorre, produzindo assim apenas uma saída (ou saídas múltiplas, dependendo do caso).

O modo como os neurônios se comportam ao longo do tempo é um processo tanto quanto como interessante. Quando uma entrada é dada a uma rede neural, a saída não aparece imediatamente porque leva um período finito para os sinais atravessarem toda a rede de neurônios, este tempo dependerá de quantos sinais de entrada existem e quantos neurônios estão conectados na rede.

Para que as redes neurais aprendam, o peso associado a cada conexão (equivalente a uma sinapse no cérebro humano) pode ser alterado em resposta a conjuntos específicos de entradas e de eventos (COPPIN, 2015). Como mencionado anteriormente, conexões que levam a respostas incorretas tendem a serem enfraquecido(a)s, enquanto os que levam a respostas corretas tendem a serem potencializado(a)s.

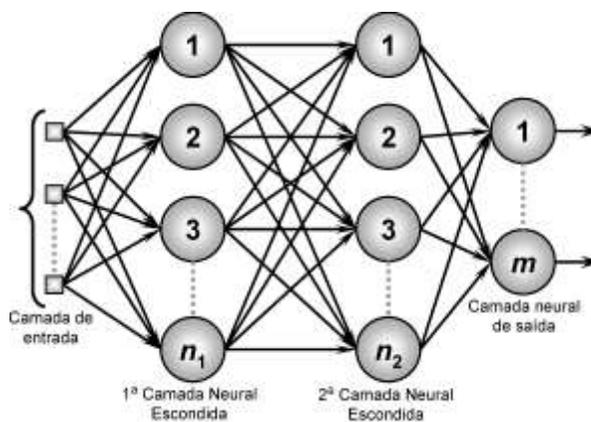
4.3.3 REDES NEURAI EM MULTICAMADAS

A maioria dos problemas do mundo real não é linearmente separável e não é encontrado resposta alguma utilizando apenas uma rede neural simples (pelo menos não para problemas de maior complexidade, que vá além de classificar entre duas opções um resultado), de apenas uma camada. Embora o modelo de uma simples rede neural seja interessante para estudar o modo como os neurônios artificiais podem funcionar, é mais interessante ainda fazer uma combinação entre estas camadas para criar algo maior com maior capacidade.

Como já foi dito, redes neurais consistem em um número finito de neurônios conectados, geralmente organizado em camadas. Quando se combina mais camadas a estes neurônios, é criada uma rede neural multicamada, que é capaz de modelar uma função que não seja linearmente separável.

Uma arquitetura típica de uma rede neural multicamada é mostrada na Figura 13.

Figura 13: Estrutura típica de uma rede neural multicamada



Fonte: Ilustração por Sandro Moreira

A primeira camada é a camada de entrada. Cada nó (ou neurônio) nesta camada recebe um sinal único de entrada. Na verdade, geralmente os nós desta camada não são neurônios, mas apenas atuam na passagem dos sinais de entrada para os nós da próxima camada, que é neste caso uma camada oculta.

Uma rede neural pode possuir uma ou mais camadas ocultas que contenham os neurônios que realmente fazem o trabalho. Observa-se que cada sinal de entrada é transmitido a cada um dos nós nesta camada e que a saída de cada nó nesta camada é transmitida a cada nó da camada final, que é a camada de saída. A camada de saída opera a última etapa de processamento e envia sinais de saída.

Quando os dados são alimentados a partir dos nós de entrada em direção aos nós de saída, caracteriza-se uma rede neural recorrente, nas quais alguns dados são passados de volta, a partir dos nós de saída em direção aos nós de entrada.

Tipicamente, uma rede neural recorrente consiste em uma camada de entrada, uma ou duas camadas ocultas e uma camada de saída e podem possuir entre 10 a 1000 neurônios em cada camada (COPPIN, 2015).

4.3.4 AJUSTE DE PESOS EM REDES NEURAIIS MULTICAMADAS

Redes neurais em multicamadas aprendem do mesmo jeito que redes neurais com apenas uma camada, a diferença principal é que em uma rede com multicamadas, cada neurônio tem pesos associados as suas entradas e há muito mais pesos a serem ajustados quando um erro é cometido para um fragmento de dados de treinamento. Um método que é utilizado normalmente é o método da retropropagação.

4.3.4.1 MÉTODO DA RETROPROPAGAÇÃO

Em vez de usar uma simples função degrau, como ilustrado no exemplo de uma rede neural com apenas uma camada, redes neurais com multicamadas com retropropagação geralmente usam a função sigmoide (Equação 4).

Equação 4: Função Sigmoide

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Assim como em redes neurais simples, o algoritmo de retropropagação começa atribuindo valores aleatórios aos pesos na rede, que geralmente recebem pequenos valores na faixa de -0.5 e +0.5. Como alternativa, os pesos podem ser normalmente distribuídos entre $-2.4/n$ e $2.4/n$, onde n é o número de entradas da camada de entrada (FAUSETT. L, 1994).

Cada iteração do algoritmo envolve, de início, a alimentação de dados por meio da rede, desde as entradas até as saídas, passando pelas camadas ocultas. A próxima fase, que dá origem ao nome do algoritmo, envolve a alimentação de erros de volta, das saídas às entradas. Estes valores de erro são alimentados de volta, através da rede, fazendo mudanças nos pesos dos nós ao longo do caminho. O algoritmo segue deste modo, até que as saídas produzidas para os dados de treinamento aproximem-se suficientemente dos valores desejados – em outras palavras, até que os valores de erro sejam suficientemente pequenos (COPPIN, 2015).

Como a função sigmoide não pode realmente chegar a 0 ou a 1, pois se trata de uma função dependente de uma exponencial negativa, é comum aceitar valores como 0.9 representando 1 e 0.1 representando 0.

Considerando uma rede neural de três camadas e utilizando i para representar os nós da camada de entrada, j para representar nós na camada oculta e k para representar nós na camada de saída. Então, por exemplo, w_{ij} se refere ao peso de uma conexão entre um nó na camada de entrada e um nó na camada oculta. A função que é usada para gerar o valor de saída de um nó j na rede pode ser vista na Equação 5.

Equação 5: Gerador de valor de saída de um nó da camada oculta

$$X_j = \sum_{i=1}^n x_i * w_{ij} - \theta_j ; Y_j = \frac{1}{1 + e^{-X_j}}$$

Sendo n o número de entradas para o nó j , w_{ij} o peso da conexão entre cada nó i e cada nó j , que é estabelecido em um valor aleatório entre 0 e 1, x_i é o valor de entrada para o nó de entrada i e Y_j é o valor de saída produzido pelo nó j .

Uma vez que as entradas tenham sido alimentadas por meio da rede para produzirem saídas, um gradiente de erro é calculado para cada nó k na camada de saída.

O sinal de erro para k é definido pela diferença entre o valor desejado e o valor atual para aquele nó (isto no caso do método supervisionado).

Equação 6: Erro

$$e_k = d_k - y_k$$

Onde d_k é o valor desejado para o nó k e y_k é o valor atual desta iteração. O gradiente de erro para o nó de saída k é definido como o valor de erro para este nó, multiplicado pela derivada da função ativação (Equação 7).

Equação 7: Gradiente de Erro na forma diferencial

$$\delta_k = \frac{\partial y_k}{\partial x_k} e_k$$

Onde x_k é a soma ponderada dos valores de entrada do nó k . Como y é definido como uma função sigmoide de x , pode-se diferenciar a função sigmoide para obter o gradiente de erro de forma mais simples (Equação 8).

Equação 8: Gradiente de Erro

$$\delta_k = y_k(1 - y_k)e_k$$

De maneira semelhante, pode-se calcular um gradiente de erro para cada nó j na camada oculta utilizando a Equação 9.

Equação 9: Gradiente de Erro para cada nó j

$$\delta_j = y_j(1 - y_j) \sum_{k=1}^n w_{jk} \delta_k$$

Onde n é o número de nós na camada de saída, ou seja, o número de saídas de cada nó na camada oculta. Desta forma, cada peso na rede, w_{ij} ou w_{jk} , é atualizado de acordo com as Equações 10 e 11.

Equação 10: Atualização do peso entre camada de entrada e camada oculta

$$w'_{ij} = w_{ij} + \alpha x_i \delta_j$$

Equação 11: Atualização do peso entre camada oculta e camada de saída

$$w'_{jk} = w_{jk} + \alpha y_j \delta_k$$

Onde x_i é o valor de entrada para o nó i e α é a taxa de aprendizado, que é um número positivo abaixo de 1 e que não deve ser muito alto, para garantir uma melhor performance do aprendizado (COPPIN, 2015).

Este método é conhecido por gradiente descendente, pois ele envolve descer a superfície que representa a função de erro pelo caminho mais íngreme (uma vez que se trata de um gradiente), para tentar encontrar o mínimo espaço de erro, o que representará o conjunto de pesos que fornece o melhor desempenho da rede.

Na verdade, a iteração do algoritmo de retropropagação é geralmente concluída quando a soma dos quadrados dos erros dos valores de saída, para todos os dados de treinamento de uma época, for menor do que um limiar t , tal como 0.001. Em outras palavras, os pesos vão se alterando favorecendo os neurônios cuja multiplicação leva a resposta correta para o problema até que o coeficiente quadrático de regressão (R^2) se aproxime de 1 com pelo menos, 3 casas decimais de precisão.

Com este método, os nós ocultos não possuem um valor de erro certo por não haver um valor de saída específico para estes nós. Neste caso, o peso de cada conexão entre um nó da camada oculta e um nó de saída é multiplicado pelo erro daquele nó de saída para tentar distribuir a responsabilidade entre os nós na camada oculta de acordo com quanto cada um contribui para o erro.

Uma observação importante a se fazer é que este método de retropropagação não ocorre na natureza, isto é, no cérebro humano. Além disto, ela é bastante ineficiente e tende a ser muito lenta para uso na solução de problemas reais. Com alguns problemas mais simples, podem-se levar centenas, ou até mesmo milhares de épocas para alcançar um nível satisfatoriamente baixo de erro (COPPIN, 2015).

Um método comum usado para melhorar o desempenho da retropropagação é incluir o *momentum* na fórmula que é usada para modificar os pesos a cada iteração.

O momentum leva em consideração a extensão da mudança de um peso específico na iteração anterior. Usa-se t para representar a iteração atual e $t-1$ para representar a iteração anterior. Então, pode-se escrever as Equações 10 e 11 na forma das Equações 12 e 13.

Equação 12: Atualização do peso entre camada de entrada e camada oculta melhorada

$$\Delta w_{ij}(t) = \alpha x_i \delta_j + \beta \Delta w_{ij}(t-1)$$

Equação 13: Atualização do peso entre camada oculta e camada de saída melhorada

$$\Delta w_{jk} = \alpha y_i \delta_k + \beta \Delta w_{jk}(t-1)$$

$\Delta w_{ij}(t)$ é a quantidade que é adicionada ao peso da conexão entre os nós i e j , w_{ij} , na iteração t . β é o valor do momentum, que é um número positivo entre 0 e 1. Tipicamente, é usado um valor razoavelmente alto, como 0.95 (COPPIN, 2015).

Esta regra também pode ser encontrada na literatura como regra delta generalizada.

A inclusão do valor de momentum tem como benefício de fazer com que o método de retropropagação evite mínimos locais e também se mova com maior velocidade nas áreas onde o espaço de erro não esteja mudando.

Um método alternativo para amplificar o desempenho da retropropagação é usar a função de ativação da tangente hiperbólica em vez de usar a função sigmoide, o que tende a permitir que a convergência da rede por uma solução ocorra em menos iterações. A função tangente hiperbólica é definida pela Equação 14.

$$\tanh(x) = \frac{2a}{1 + e^{-bx}} - a$$

Onde a e b são constantes.

Um último meio para melhorar o desempenho da retropropagação é variar o valor da taxa de aprendizado, α , durante a fase de treinamento da rede. Duas heurísticas propostas por R. A. Jacobs (1988) usam a direção de mudança (aumento ou decréscimo) da soma dos quadrados dos erros de uma época para a próxima, para determinar a mudança na taxa de aprendizado (COPPIN, 2015):

- Se, para muitas épocas, a soma dos quadrados dos erros mudar na mesma direção, a taxa de aprendizado deve ser aumentada.
- Se a soma dos quadrados dos erros alterar na direção nas suas mudanças por muitas épocas, a taxa de aprendizado deve ser reduzida.

Usando estas duas heurísticas, em combinação com a regra delta generalizada, o desempenho do algoritmo de retropropagação é melhorado de forma significativamente.

4.3.5 REDES NEURAIIS RECORRENTES

As redes estudadas até o momento são redes de alimentação adiante, e estas são acíclicas, no sentido que não há ciclos na rede, porque os dados passam das entradas às saídas, mas o contrário não ocorre. Uma vez que uma rede de alimentação adiante tenha sido treinada, seu estado é fixado e não se altera conforme novos dados sejam apresentados a ela. Em outras palavras, estes tipos de redes não possuem a capacidade de armazenar dados, apenas de processá-los.

Uma rede neural recorrente pode ter conexões que voltem dos nós de saída aos nós de entrada e, na verdade, pode ter conexões arbitrárias entre quaisquer nós. Desta forma, o estado interno de uma rede recorrente pode ser alterado conforme conjuntos de entradas sejam apresentados a ela e, assim, pode-se dizer que ela possui um tipo de memória.

Desta maneira, pode-se utilizar este tipo de rede para solucionar problemas dos quais não dependem apenas das entradas usuais dadas pelo usuário, mas sim, utiliza-se as saídas de uma época para serem os dados de entrada da próxima.

Um exemplo, estas redes poderiam serem utilizadas para prever o preço de mercado de uma específica ação, baseado em todos os valores prévios ou até mesmo poderia fazer uma previsão de tempo para o dia seguinte, baseando-se em como o tempo tem estado ultimamente. Claro que estes tipos de previsões não são totalmente confiáveis, pois isso apenas se trata de uma análise estatística dos dados, e como se sabe o preço de mercado de uma ação não varia no tempo através de um modelo matemático, sendo totalmente imprevisível e também não seria aplicado para previsão de tempo, uma vez que este

depende de um modelo físico que contém todas as leis conhecidas da natureza para realizar tal previsão. Agora, pode-se citar um lugar de boa aplicação para determinar marés em costas do país, uma vez que este fenômeno é cíclico e totalmente previsível. Em outras palavras, estas redes podem ser usadas para descrever um modelo matemático de fenômenos cíclicos, com isto, pode-se realizar previsões, do contrário, para fenômenos acíclicos, isto é, que não possuem variabilidade visível, as redes neurais recorrentes apenas seria útil para uma análise estatística dos dados – como é o caso deste projeto.

Obviamente, devido à ausência de memória, redes de alimentação adiante não são capazes de executar tais tarefas.

Ao aprender, a rede neural recorrente alimenta suas entradas por meio da rede, incluindo alimentação de dados de volta, das saídas às entradas, repetindo este processo até que os valores das saídas não mudem mais. Neste ponto, diz-se que a rede está em um equilíbrio ou estabilidade.

Pode-se considerar que uma rede neural recorrente é uma memória que é capaz de aprender um conjunto de estados. Uma vez que tal rede tenha sido treinada, para qualquer entrada dada ela terá como saída o atrator que esteja mais próximo da entrada fornecida.

5.0 METODOLOGIA

A primeira tentativa para determinar atrasos de sinais vindo de componentes de lentes gravitacionais aconteceu em 1984 (Florentin-Nielsen,1984) utilizando um único telescópio. O atraso calculado foi de aproximadamente 1,5 anos para o quasar 0957+561 – primeira lente gravitacional descoberta. Porém os dados estavam muito ruidosos para se levar em conta a determinação daquele atraso. No mesmo ano outra curva de luz foi publicada (Schild e Weekes, 1984) onde foi notada uma clara indicação de variabilidade. Em 1986 foi estimado um atraso de aproximadamente 1,03 anos (Schild e Cholfin,1986). Outros trabalhos mostraram um atraso de cerca de 415 dias ou seja cerca de 1,14 anos.

Segundo Falco, Gorenstein e Shapiro, 1990 a relação entre o atraso de sinais vindo de duas componentes A e B do quasar 0957+561 (exemplo deste atraso: ~1 ano), a velocidade de dispersão da galáxia-lente (exemplo: ~390 km/s) e a constante de Hubble (H_0) é dada pela Equação 15.

Equação 15: Constante de Hubble

$$H_0 = \frac{V_d (90 \pm 20)}{\Delta T}$$

Onde “ V_d ” é a velocidade de dispersão da galáxia-lente e “ ΔT ” é o atraso no tempo das componentes. A dimensão de H_0 é Km/s/Mpc.

Dessa forma sabendo-se o atraso entre os sinais A e B pode-se calcular a constante de Hubble. Para isso um quasar tem de ser observado em um período longo de tempo (cerca de 10 anos) para que a variabilidade das componentes possa ser registrada e o atraso possa ser calculado.

Sabendo-se que o inverso da constante de Hubble tem dimensão de tempo, pode-se calcular a idade do Universo. Considerando-se que galáxias se afastam uma das outras com velocidades constantes “ v ”, supõe-se que elas estavam juntas num tempo “ d/v ” no passado, o que pode ser traduzido pela Equação 16.

Equação 16: Idade do Universo

$$Idade\ do\ Universo = \frac{1}{H_0}$$

Na verdade, a consideração de velocidade constante não é a correta e somente foi usada neste projeto para ter uma estimativa da idade do universo a partir do conhecimento da constante de Hubble.

Foi realizada uma pesquisa sobre os quasares que poderiam ser candidatos à Lentes Gravitacionais e foram encontrados cinco possíveis sistemas já citados no Capítulo 3. Os dados para o desenvolvimento da pesquisa foram obtidos através do Centro de Dados Astronômicos de Estrasburgo (CDS). Com os dados em mãos, pode-se plotar a magnitude relativa de cada componente das lentes gravitacionais em função dos dias para verificar se há variabilidades; caso haja, observar possíveis ocorrências de explosões na fonte. Desta comparação entre os objetos foram extraídas as diferenças de tempo entre as estruturas observadas nas duas ou mais curvas de luz.

Inicialmente, planejava-se a realização de observações no rádio observatório de Itapetinga mensalmente, para se obter a densidade de fluxo de certas fontes utilizando a técnica de varreduras, com durações de 20,0 segundos e cada observação com 30 varreduras, porém, a antena estava em manutenção, e logo após sua reabertura ocorreu a pandemia devido ao COVID-19.

Para a análise das curvas de Luz o Software utilizado foi o MATLAB, onde foram plotados os dados obtidos do Observatório de Estrasburgo e analisados utilizando métodos de análises matemáticas e estatísticas, como regressão, interpolação, interpolador harmônico de Fourier e finalmente, a utilização de redes neurais para a obtenção de pontos faltantes na curva analisando o padrão da variabilidade dos candidatos a lentes gravitacionais.

5.1 MÉTODO DE ANÁLISE DAS CURVAS DE LUZ

Através dos dados obtidos do Observatório de Estrasburgo, pode-se analisar a variabilidade dos candidatos e observar o comportamento das curvas de luz em relação às explosões que ocorrem nas fontes, eventos nos quais são de extrema importância para determinar o *time delay* entre as imagens oriundas do efeito de lente gravitacional.

Inicialmente, foi utilizado o método do Interpolador Harmônico de Fourier, que se baseia em uma função dada em termos de senos e cossenos. Embora utilizando este método obteve-se resultados muito acurados, porém pouco precisos. A incerteza associada aos valores é de ordem significativa para a análise da constante de Hubble e a idade do Universo. Foi utilizada uma técnica matemática com aplicação da Série de Fourier. Isto é, para cada componente de cada quasar candidato à lente gravitacional foi ajustado uma função harmônica para os conjuntos de dados buscando-se o maior coeficiente quadrático de regressão (R^2) possível, uma vez que este coeficiente informa o percentual de erro do polinômio ajustado que é removido utilizando o ajuste escolhido.

Posteriormente, foi construída uma rede neural buscando os pontos faltantes a curva reconhecendo o padrão nos dados. Usualmente, é conveniente visualizar os neurônios como camadas arranjadas, onde a chave do processo é a função de ativação do mesmo, contendo o mesmo padrão nos pesos entre as conexões dos neurônios anteriores e subsequentes.

O arranjo dos neurônios em camadas e suas conexões com outros neurônios de outras camadas é chamada de *arquitetura neural*. Os dados de entrada foram os pontos da curva de luz, e entre um ponto e outro foi utilizado o método de *spline cúbico*, para obter as *unidades escondidas* das conexões, que servem de auxiliares para realizar o ajuste de pesos para cada época de treinamento.

O método de ajuste dos valores de peso (treinamento) é uma importante característica de uma rede neural. Por conveniência, há dois tipos diferentes de treinamento: o supervisionado e não supervisionado. Muitas tarefas da rede neural precisam ser treinadas para realizarem uma melhor performance em cada época, melhorando sua capacidade de encontrar soluções corretas e enfraquecendo a obtenção de soluções incorretas. Quando é utilizado o método supervisionado para o ajuste dos valores de peso, o valor de saída teórico da rede normalmente é conhecido e toda vez que a rede é ativada, o valor de saída prático da rede é comparado com o valor teórico, e utilizando a fórmula de treinamento de perceptron (um simples neurônio proposto por Rosenblatt (1958) usado para classificar suas entradas em uma de duas categorias) dada pela Equação 17.

Equação 17

$$w'_i = w_i + (a * x_i * e)$$

Onde a é a taxa de aprendizagem, x_i o valor de entrada e e o erro entre o valor de saída da rede e o valor esperado. Uma vez que esta modificação dos pesos tenha ocorrido, o próximo fragmento de dados de treinamento será empregado do mesmo modo. Uma vez

que todos os dados de treinamento tenham sido aplicados, o processo é iniciado novamente, até que todos os pesos estejam corretos e todos os erros sejam zerados. Cada iteração deste processo é conhecida como uma época.

O treinamento não supervisionado consiste em uma auto-organização da rede neural em grupos similares com o objetivo de treinar os dados para um valor cuja natureza seja similar em características de um certo grupo de valores oriundos dos dados iniciais. Uma sequência de vetores é providenciada, mas não há um vetor “alvo” específico. A rede modifica os pesos para que o vetor mais provável de saída.

Nesta metodologia, foi utilizado o treinamento supervisionado, uma vez que os dados de saída são estimados através da metodologia da interpolação de Fourier. Portanto, foi realizado um trabalho mútuo entre o aprendizado de máquina para aumentar a precisão dos resultados obtidos pelo método da interpolação de Fourier.

Para fins de comparação e determinação da diferença de tempo das curvas de luz, foram plotadas as componentes em um só gráfico e conhecendo-se o polinômio ajustado pelo método de Fourier, pode-se determinar os valores dos pontos de máxima magnitude, ou seja, possíveis explosões na fonte. Comparando-se a defasagem de dias entre a explosão da componente A da fonte com a componente B, e com as demais componentes caso haja, pode-se assim determinar a defasagem de tempo entre os picos de magnitude.

Posteriormente, utilizando os dados e os ajustes realizados, foi levado ao algoritmo produzido para a classificação do ajuste e encontrar pontos que complementam a curva sem a necessidade de outro ajuste, com isto, pode-se determinar o *time delay* entre as fontes com uma precisão superior do método de Fourier.

O valor da velocidade de dispersão de cada galáxia-lente foi obtido através de pesquisas em referências na literatura.

6.0 CURVAS DE LUZ ANALISADAS

Utilizando os dados obtidos através das observações do Centro de Dados Astronômicos de Estrasburgo (CDS), pode-se plotar a magnitude relativa de cada fonte em função dos dias de observação e aplicar os métodos descritos na Metodologia, sendo o primeiro uma aplicação da interpolação harmônica de Fourier e a segunda utilizando um algoritmo feito com base em redes neurais recorrentes.

Dos Gráficos 1, 9, 15, 23 e 29 pode-se observar o comportamento da luz oriunda destes objetos e suas componentes e também é possível notar que há picos de magnitude, ou seja, prováveis explosões nas fontes.

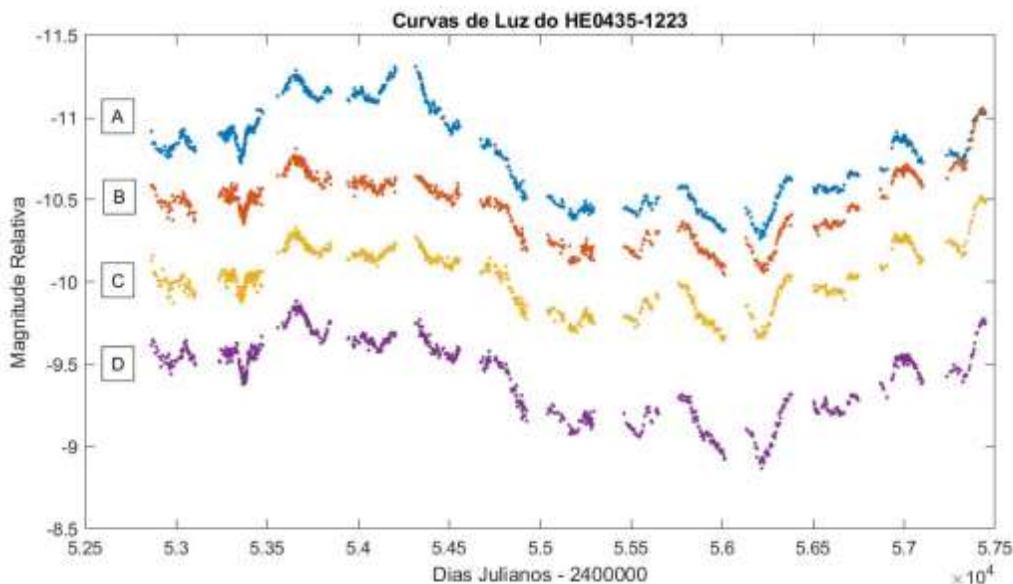
No eixo das abcissas utilizou-se uma técnica onde se subtrai 2400000 do dia Juliano, caso contrário, não seria uma boa maneira de apresentar os dados no eixo, isto é, o início das observações do Observatório de Estrasburgo começou no dia juliano $52864+2400000$, totalizando 2452864, que no calendário gregoriano utilizado normalmente é o dia 12 de agosto de 2003 para a lente gravitacional HE0435-1223.

No total, foram 13 anos de observações de cada candidato a lente que estão nos Gráficos 1, 9, 15, 23 e 29, de 2003 a 2016, devido a isto pode-se observar uma variabilidade periódica nos dados e conseqüentemente, adotar métodos de previsão destes dados para estimar os valores dos pontos faltantes.

Na próxima seção serão apresentados os gráficos para cada candidato à lente gravitacional juntamente com suas componentes e utilizando-se a série harmônica de Fourier e também utilizando redes neurais recorrentes, pode-se fazer uma previsão nos intervalos das componentes, determinando assim um ajuste razoável para cada conjunto de dados e a criação de pontos para poder determinar o *time delay*.

6.1 HE0435-1223

Gráfico 1: Dados do HE0435-1223



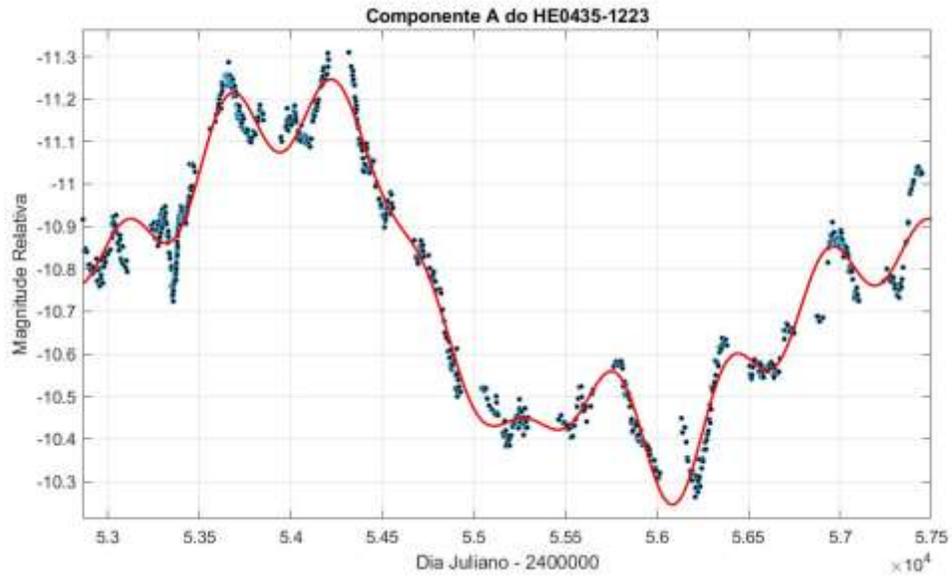
Fonte: Acervo do Autor

Utilizando o primeiro método para a análise das curvas utilizando a interpolação harmônica de Fourier, pode-se determinar os ajustes para cada componente da lente gravitacional HE0435-1223, sendo elas as componentes A,B,C e D. Os ajustes obtidos estão nos Gráficos 2, 3, 4 e 5.

Nota-se também uma variabilidade de aparentemente 5 anos entre o máximo e o mínimo da fonte, levando a conclusão que a partir destes dados, há uma periodicidade de magnitude máxima (em módulo) de 5 anos.

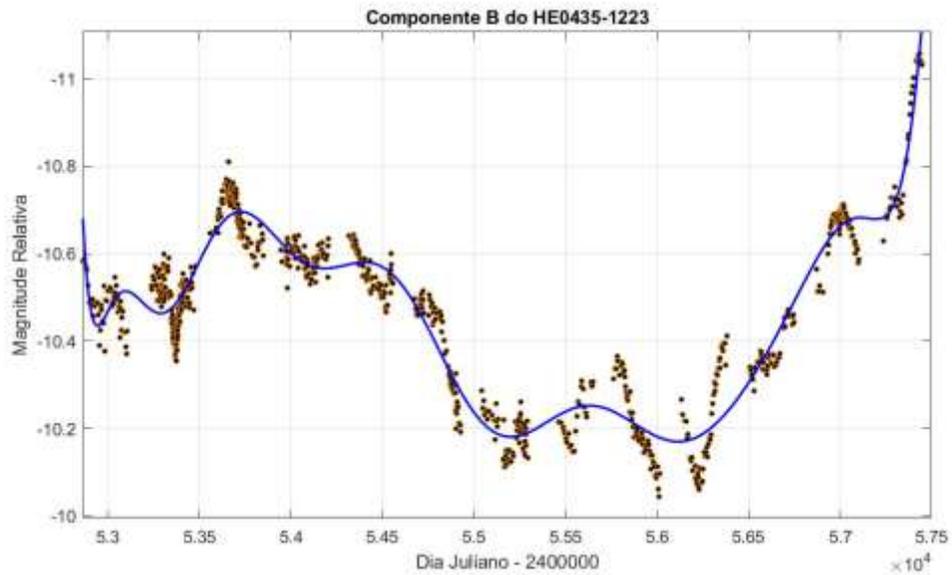
Para a componente A do HE0435-1223, o ajuste teve o seu coeficiente de regressão quadrático, isto é, o valor de R^2 de 0.9721, ou seja, cerca de 97,21% dos erros do ajuste em relação aos pontos originais foram eliminados utilizando este método para esta componente. Portanto, este é um ajuste razoável para este conjunto de dados.

Gráfico 22: Componente A do HE0435-1223



Fonte: Acervo do Autor

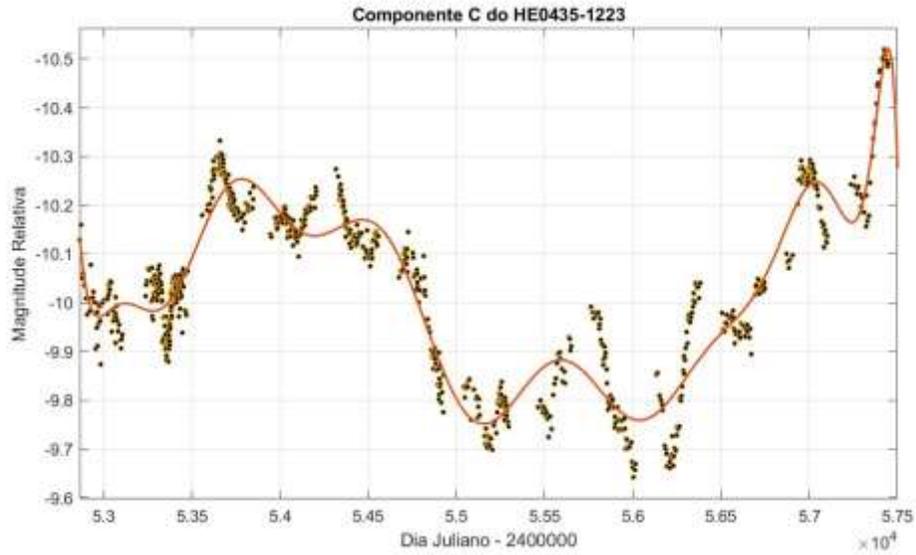
Gráfico 3: Componente B do HE0435-1223



Fonte: Acervo do Autor

Já a componente B apresentou um valor para este coeficiente um pouco menor, de 0.9244, embora não seja um valor muito baixo, este será desconsiderado para evitar maiores erros.

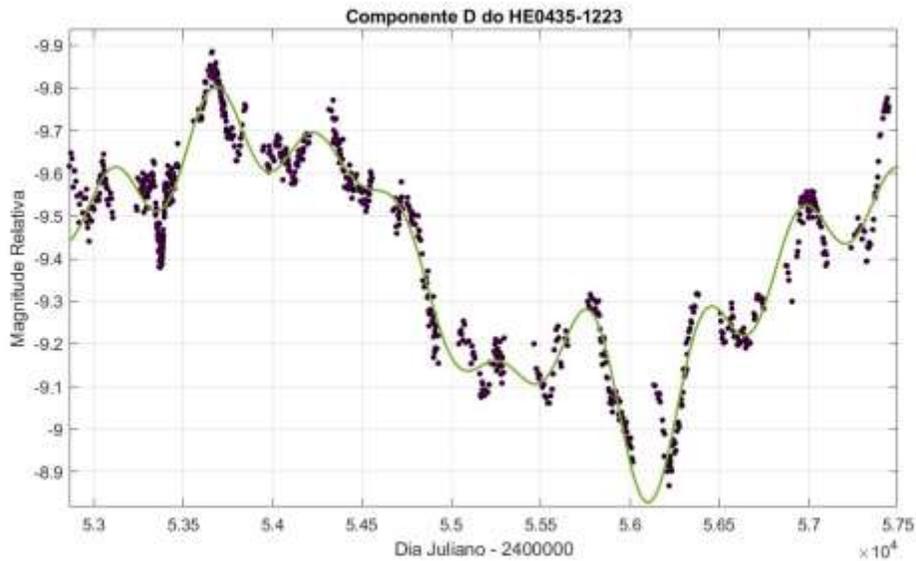
Gráfico 4: Componente C do HE0435-1223



Fonte: Acervo do Autor

A componente C apresentou um valor de 0.8996, um valor não muito bom para se considerar como um ajuste confiável.

Gráfico 5: Componente D do HE0435-1223

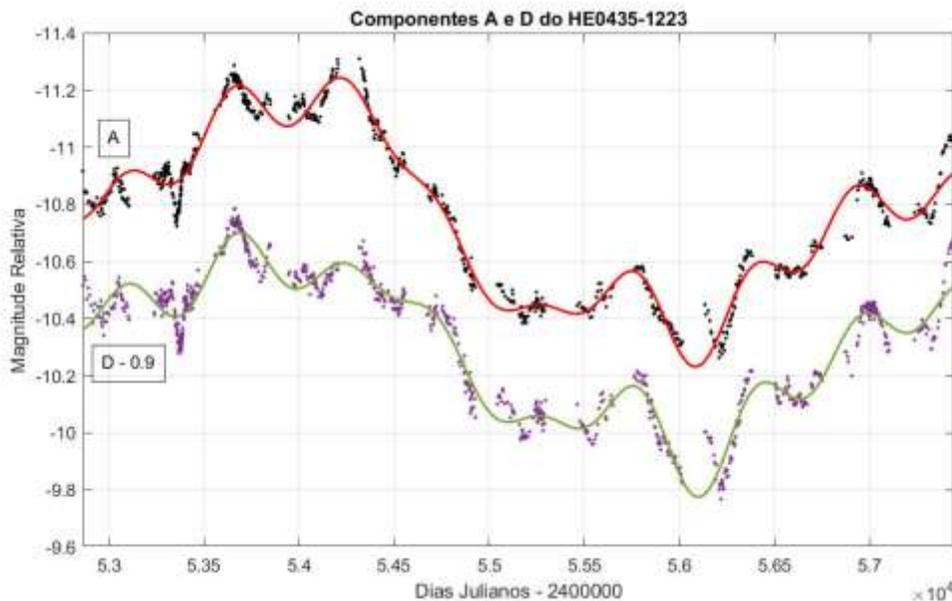


Fonte: Acervo do Autor

Finalmente, para a componente D foi calculado o valor de R^2 igual à 0.9465. Logo, pode-se concluir que os ajustes em geral são bons, porém para fins de comparação para a determinação da defasagem de tempo entre os picos das fontes será levada em consideração

a componente A e D do candidato a lente gravitacional HE0435-1223, uma vez que o ajuste para os mesmos apresentaram serem os melhores em relação ao B e o C.

Gráfico 6: Componentes A e D do candidato a lente gravitacional HE0435-1223



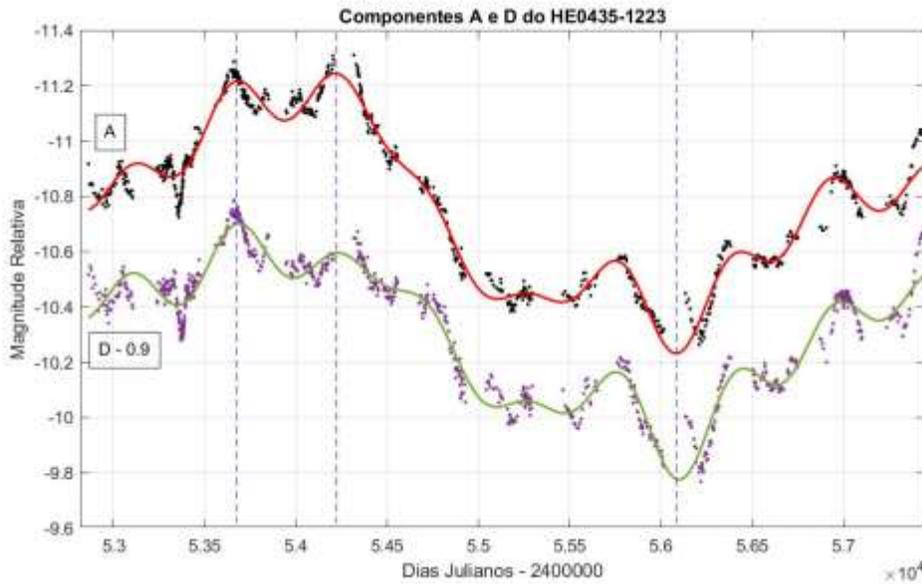
Fonte: Acervo do Autor

O próximo passo é utilizar o software para marcar os picos de máximo e de mínima magnitude relativa da curva de luz ajustada e posteriormente utilizando o polinômio ajustado, calcular o dia no qual ocorreu estes picos para cada componente. O gráfico que contém esta análise pode ser observado no Gráfico 7.

Pode-se notar que há diferença temporal entre os picos de máximo e de mínimo das curvas de luz, e para determinar a diferença de dias entre estes valores foi utilizado os polinômios ajustados pelo algoritmo e calculando-se assim a diferença entre os picos de $(8,0 \pm 0,8)$ dias. Para a determinação da velocidade de dispersão da galáxia que está sofrendo o efeito de lente, foi realizado uma pesquisa na literatura e encontrado o Redshift da mesma como $Z = (1,693 \pm 0,103)$ (H0LiCOW Paper IV).

Pode-se ainda estimar a velocidade de dispersão da galáxia através da Equação 7. Se o objeto tiver um desvio para o vermelho muito grande, uma correção relativística deve ser realizada, que no caso será o que deve ser feito para estimar a velocidade de dispersão da galáxia, uma vez que segundo o H0LiCOW Paper IV, o desvio para o vermelho da galáxia que está sob o efeito de lente gravitacional do quasar HE0435-1223 é bem maior do que 1. A correção está apresentada na Equação 19.

Gráfico 7: Determinação do time delay entre os picos de máximo e de mínimo do HE0435-1223



Fonte: Acervo do Autor

Equação 18: Estimativa do Redshift para desvios pequenos ($z < 1$)

$$z = \frac{v}{c}$$

Equação 19: Redshift para grandes desvios ($z > 1$)

$$z = \sqrt{\frac{1 + \frac{v}{c}}{1 - \frac{v}{c}}} - 1$$

Utilizando os resultados obtidos pelo ajuste e pela pesquisa na literatura em relação à velocidade de dispersão, pode-se utilizar a Equação 4 para estimar a Constante de Hubble por este método que foi realizado.

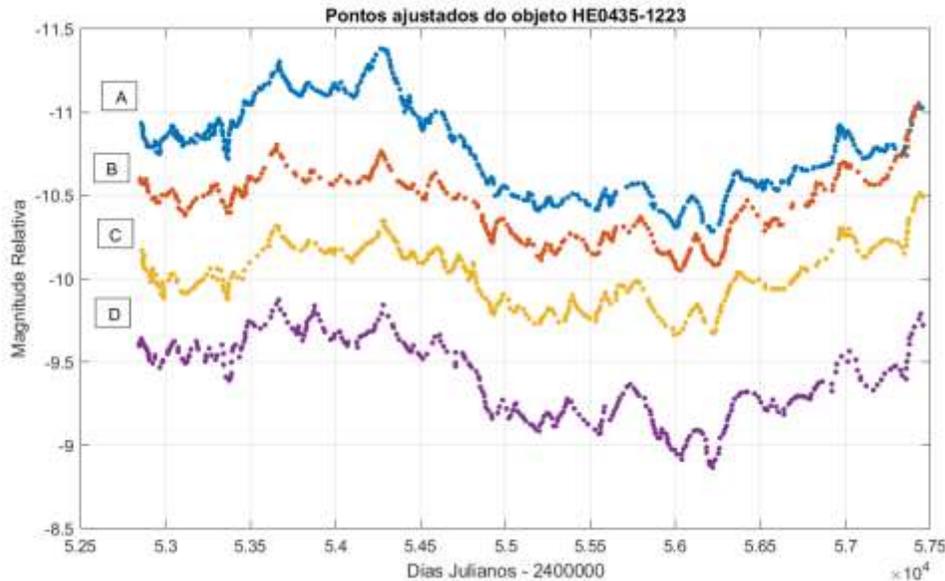
Portanto, utilizando-se o método que foi descrito e os valores obtidos, determina-se que a Constante de Hubble possui o valor de aproximadamente $H_0 = (71,6 \pm 4,7)$ Km/s/Mpc pelo método de Interpolação Harmônica de Fourier.

Consequentemente, segundo a Equação 16, pode-se estimar a idade do Universo utilizando-se a Constante de Hubble calculada através da lente gravitacional HE0435-1223, obtendo-se o valor para a Idade do Universo de $T_U = (4,18 \pm 0,02) \times 10^{17}$ segundos, isto é, a

idade do Universo estimada através dos métodos descritos é de cerca de $T_U = (1,36 \pm 0,02) \times 10^{10}$ anos, ou seja, aproximadamente 13,6 bilhões de anos, com uma incerteza de mais ou menos 0,02 bilhões de anos.

Já utilizando o segundo método, que é pelo algoritmo feito com base em redes neurais recorrentes, com aprendizado supervisionado, o Gráfico 8 foi obtido.

Gráfico 8: Pontos criados pela RNR do objeto HE0435-1223



Fonte: Acervo do Autor

Desta forma, como faltavam poucos dados na série, pode-se realizar um ajuste dos pontos com um coeficiente de regressão (R^2) próximo de 1 com uma precisão de três casas decimais.

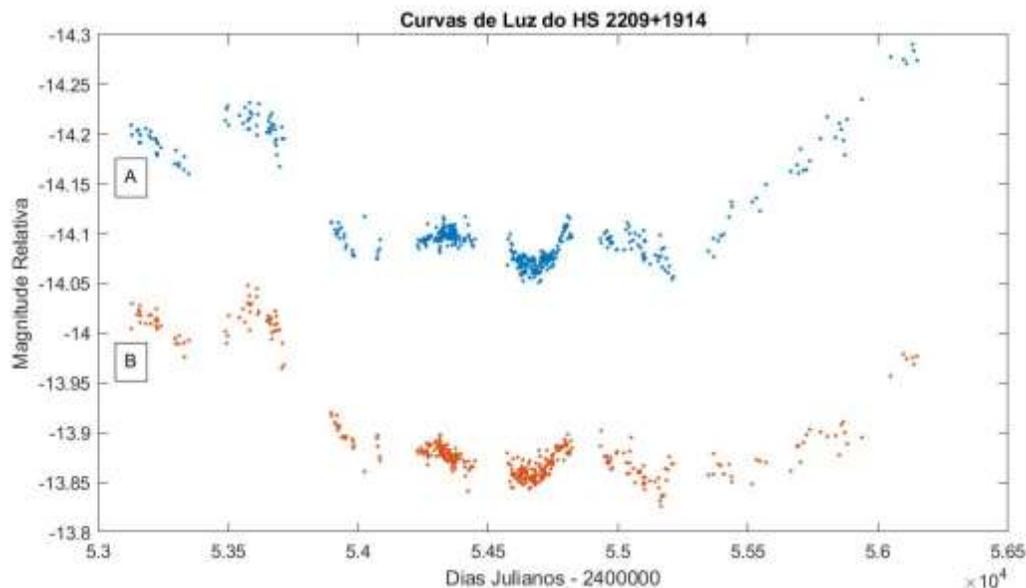
Após realizar a análise semelhante ao método descrito anteriormente, pode-se obter um *time delay* de aproximadamente $(8,32 \pm 0,47)$ dias em média de todas as curvas A, B, C e D do objeto candidato a lente gravitacional HE0435-1223.

Consequentemente, o valor calculado para a constante de Hubble foi de $H_0 = (69,4 \pm 1,7)$ Km/s/Mpc. Portanto, a idade do universo obtida pelo método das RRNs para o *time delay* do objeto HE0435-1223 foi de $T_U = (4,31 \pm 0,01) \times 10^{17}$ segundos, isto é, a idade para o universo em anos seria de $T_U = (1,36 \pm 0,01) \times 10^{10}$ anos, ou seja, 13,6 bilhões de anos, com uma incerteza associada de 0,01 bilhão para mais ou para menos.

Pode-se observar que para ambos os métodos se obteve a mesma idade para o universo, porém, com constantes de Hubble diferentes. Para resolver este problema, deve-se obter outros valores de outros candidatos e realizar um tratamento estatístico dos dados.

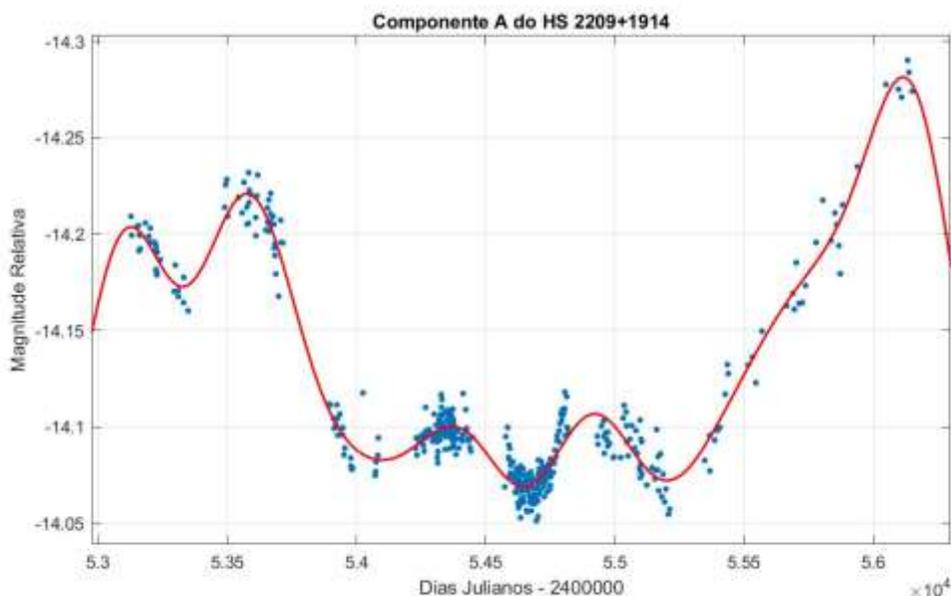
6.2 HS 2209+1914

Gráfico 9: HS 2209+1914



De maneira análoga, a análise das curvas de Luz dessa lente gravitacional será feita por cada componente, analisando o seu coeficiente quadrático de regressão (R^2) e verificando se o polinômio ajustado pelo algoritmo é razoável para descrever o conjunto de dados, caso R^2 seja um número bom, isto é, o ajuste é fiel aos dados, poderá ser feita a comparação dos ajustes determinando-se assim a defasagem de tempo entre os valores de máximo e de mínimo da função ajustada.

Gráfico 10: Componente A do HS 2209+1914

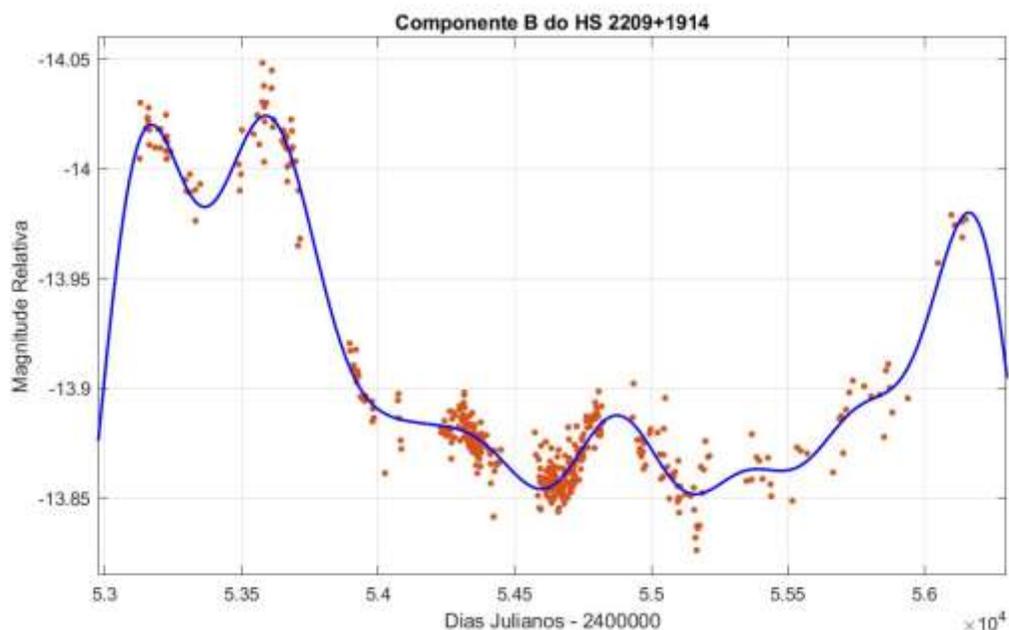


Para o ajuste da Componente A do HS 2209+1914, obteve-se o coeficiente quadrático de regressão (R^2) com um valor de 0,968, isto é, 96,80% dos erros foram eliminados da regressão utilizando-se o ajuste dado no Gráfico 10. Logo, é possível dizer que o ajuste é razoável para o conjunto de dados da componente A desta lente gravitacional.

Para o ajuste do Gráfico 11, o coeficiente quadrático de regressão foi calculado que seu valor é de 0,9735. Logo, pode-se concluir que ambos os ajustes estão razoáveis para os conjuntos de dados das componentes A e B. Portanto, deve-se fazer a comparação entre os ajustes e determinar os valores de máximo e de mínimo dos mesmos, e por fim, comparar as curvas obtendo-se a diferença de tempo entre tais valores.

No Gráfico 12 pode-se comparar as curvas de luz entre as componentes A e B desta fonte. É notável que há uma diferença de tempo entre os picos de máxima e mínima magnitude relativa, os picos de máxima podem ser prováveis explosões na fonte. Logo, é feita uma comparação destes picos e calculando-se os valores no eixo da abcissa é possível calcular a defasagem de tempo entre os mesmos.

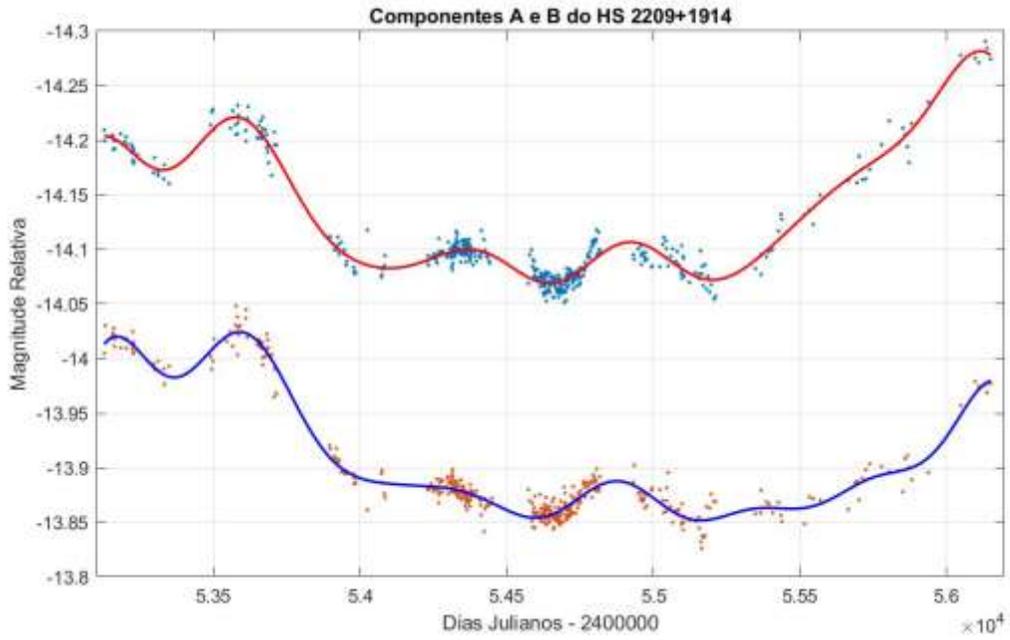
Gráfico 11: Componente B do HS 2209+1914



Fonte: Acervo do Autor

É notável que há uma defasagem entre os valores de máximo e de mínimo nestas fontes, consequentemente, utilizando-se o software pode-se determinar qual é o valor desta diferença de tempo, e o mesmo resultou em $(21,56 \pm 3,40)$ dias.

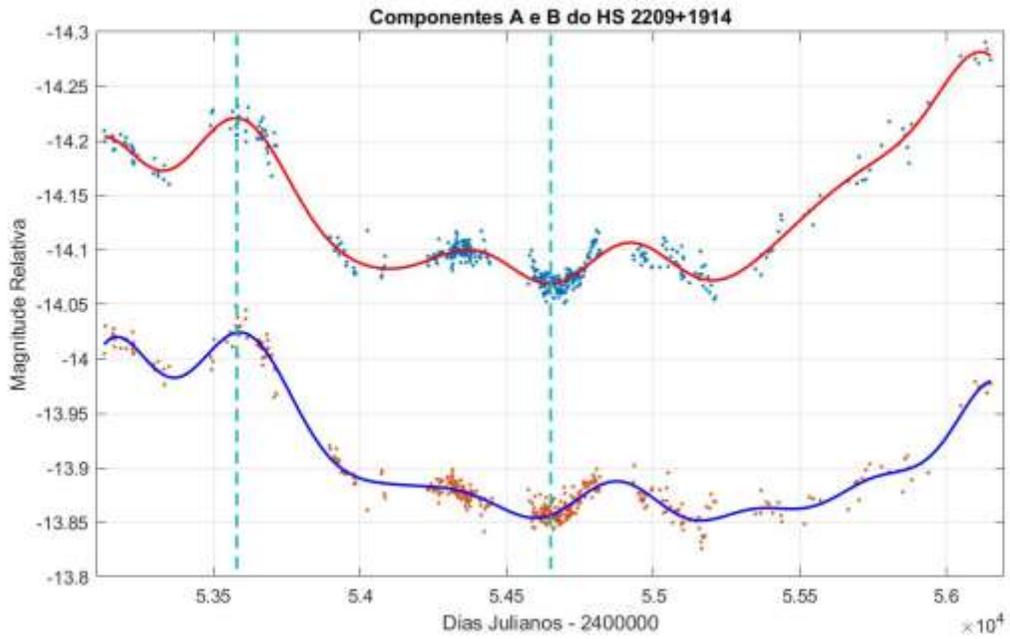
Gráfico 12: Análise das Curvas de Luz do HS 2209+1914



Fonte: Acervo do Autor

Realizando a análise descrita, tem-se o Gráfico 13.

Gráfico 13: Análise das Curvas de Luz do HS 2209+1914



Fonte: Acervo do Autor

O objeto que está sob o efeito de lente gravitacional deste quasar possui um Redshift de cerca 1,07 (*Fonte: CDS*), onde o intervalo de confiança não foi informado. Consequentemente, utilizando-se a Equação 1, pode-se estimar a constante de Hubble para esta fonte.

A constante de Hubble estimada utilizando os dados do HS 2209+1914 e o método que foi descrito é de $H_0 = (62,8 \pm 4,8)$ Km/s/Mpc.

Segundo a Equação 2, pode-se estimar a idade do Universo calculando o inverso da constante de Hubble, e este valor resultou em $T_U = (4,92 \pm 0,07) \times 10^{17}$ segundos, transformando em anos, a idade do Universo estimada por este método é de $T_U = (1,55 \pm 0,03) \times 10^{10}$ anos, ou seja, de aproximadamente 15.5 bilhões de anos.

Já utilizando o segundo método, que é pelo algoritmo feito com base em redes neurais recorrentes, com aprendizado supervisionado, os pontos ajustados estão representados no Gráfico 14.

Gráfico 14: Pontos criados pela RNR do objeto HS2209+1914

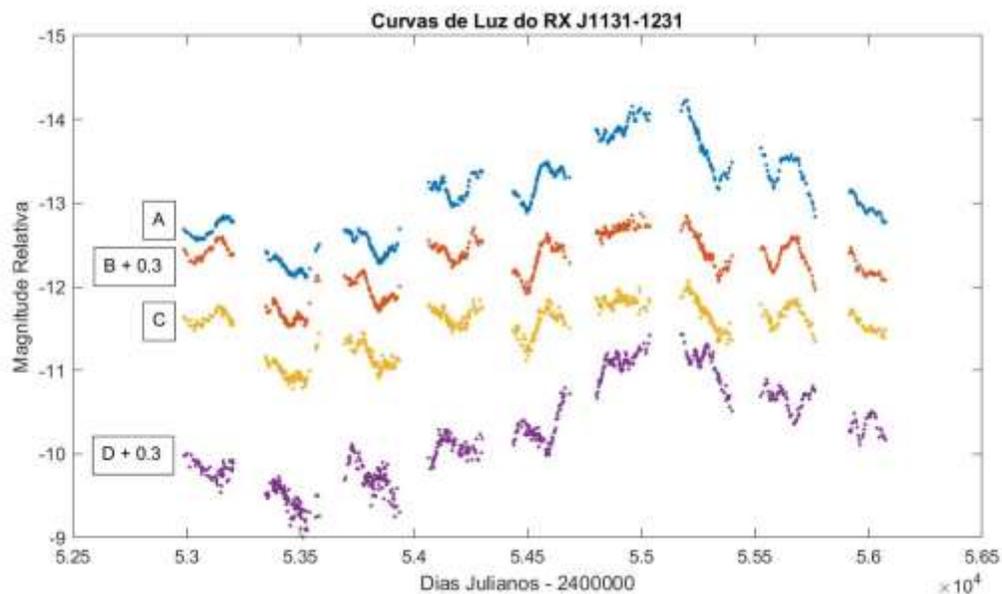


Fonte: Acervo do Autor

Através do ajuste pela RNN, pode-se determinar um *time delay* entre a componente A e B do objeto HS2209+1914 de $(19,83 \pm 1,29)$ dias. Logo, utilizando os mesmos cálculos do método anterior, foi possível obter a constante de Hubble com um valor de $H_0 = (68,3 \pm 2,4)$ Km/s/Mpc. Consequentemente o valor obtido para a idade do Universo foi de $T_U = (1,42 \pm 0,05) \times 10^{10}$ anos, ou seja, de aproximadamente 14,2 bilhões de anos com uma incerteza associada de 0,05 bilhão para mais ou para menos.

6.3 RX J1131-1231

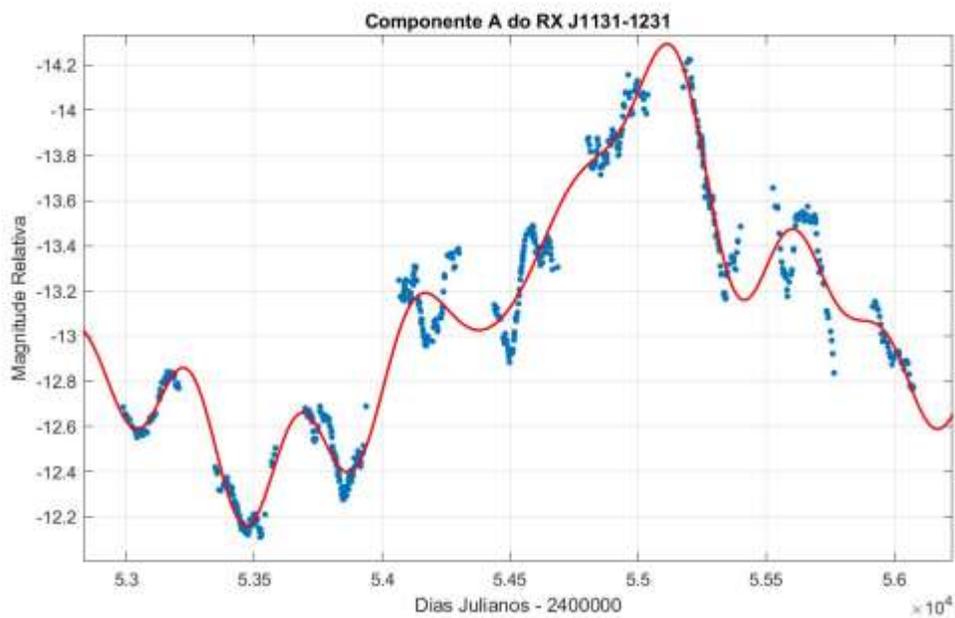
Gráfico 15: RX J1131-1231



Fonte: Acervo do Autor

Para este objeto pode-se observar as curvas de luz ajustadas das componentes A, B, C e D nos Gráficos 16,17,18 e 19 respectivamente.

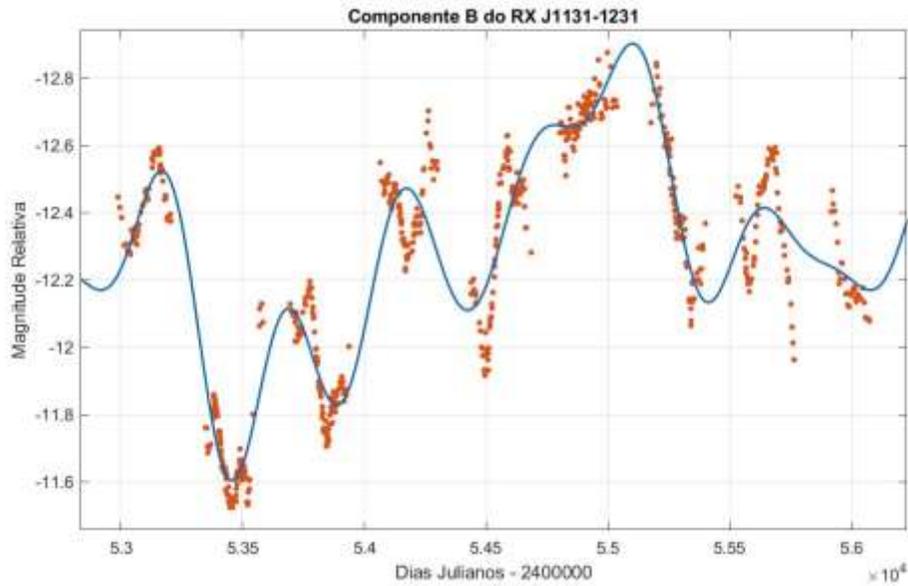
Gráfico 16: Componente A do RX J1131-1231



Fonte: Acervo do Autor

Para o ajuste do Gráfico 16, obteve-se o valor de R^2 igual a 0,9704.

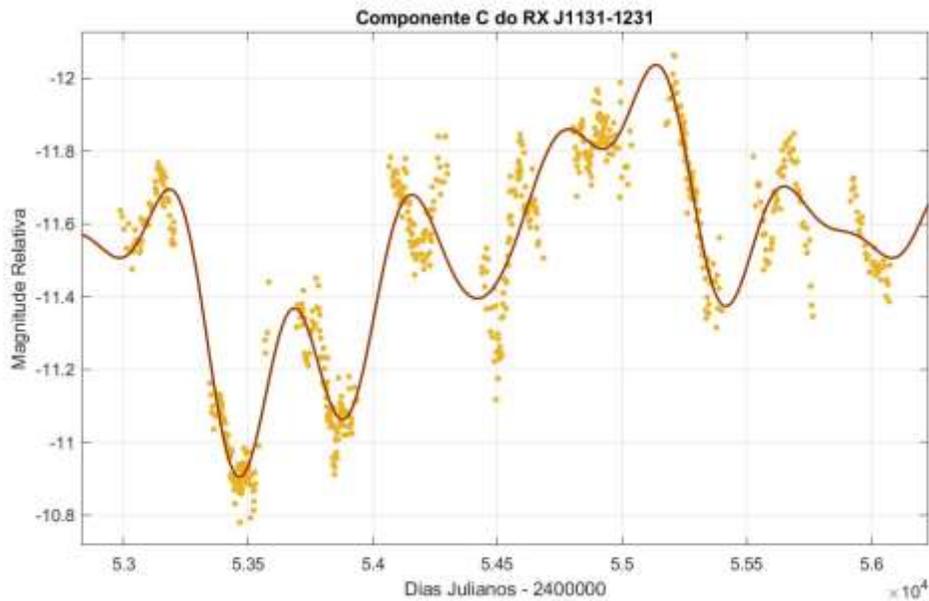
Gráfico 17: Componente B do RX J1131-1231



Fonte: Acervo do Autor

Para o ajuste do Gráfico 17, obteve-se o valor de R^2 igual a 0,9118.

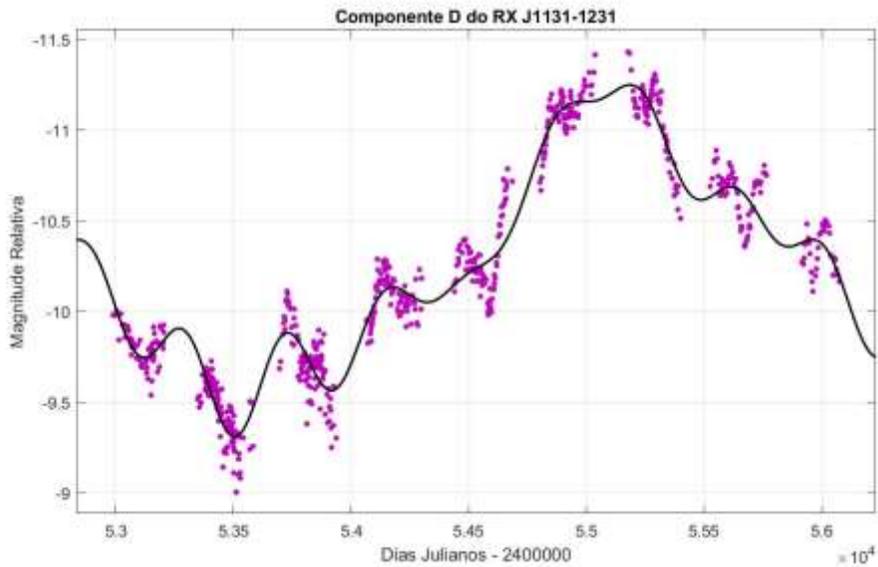
Gráfico 18: Componente C do RX J1131-1231



Fonte: Acervo do Autor

Para o ajuste na Componente C do RX J1131-1231 foi calculado o valor do R^2 igual a 0,9122.

Gráfico 19: Componente D do RX J1131-1231



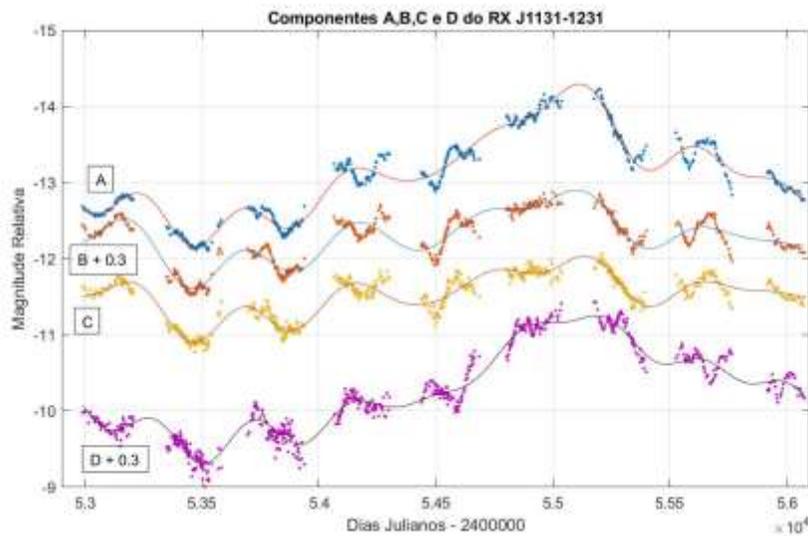
Fonte: Acervo do Autor

Para o ajuste do Gráfico 19 foi calculado o valor de R^2 igual a 0,9585.

Comparando-se os valores do coeficiente quadrático de regressão (R^2) pode-se observar que os ajustes A e D são os que possuem maiores valores, isto é, os polinômios ajustados pelo algoritmo destas componentes são os que mais se aproximam ao comportamento real dos pontos.

Porém, como todos os ajustes obtiveram um R^2 razoável, pode-se comparar todas as curvas e calcular a defasagem de tempo entre elas.

Gráfico 20: Análise das Curvas de Luz do RX J1131-1231



Fonte: Acervo do Autor

Como há quatro curvas, utilizando-se o polinômio ajustado pelo algoritmo pode-se determinar os valores dos picos e conseqüentemente a defasagem de tempo entre eles das componentes A,B,C e D. Isto é, há uma defasagem entre A e B, A e C, A e D, B e C, B e D, e C e D. Logo, foi calculado a constante de Hubble para cada comparação entre as componentes.

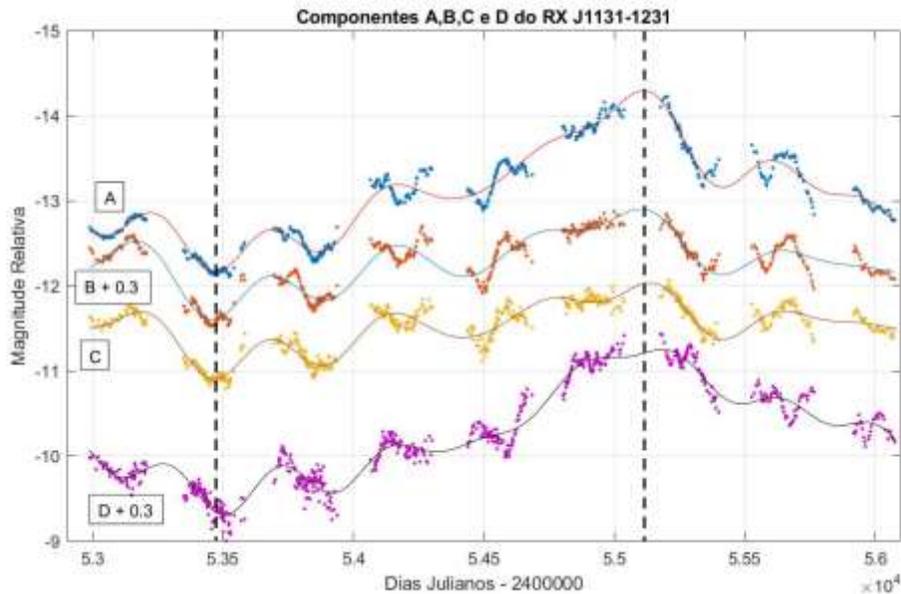
Na literatura foi encontrado que o “Redshift” para o objeto que está sob o efeito de lente gravitacional do RX J1131-1231 é de cerca de 0,654 (Suyu et al. 2014).

Logo, pode-se utilizar a Equação 15 para determinar a constante de Hubble para cada defasagem entre as componentes e a Equação 16 para estimar a idade do Universo.

Os resultados podem ser observados na Tabela 1, e no Gráfico 21 pode-se observar os valores de máximo e de mínimo traçados por uma reta que o algoritmo identificou.

Em média para esta fonte (RX J1131-1231) estimou-se que a constante de Hubble vale $H_0 = (67,4 \pm 4,2)$ Km/s/Mpc e a idade média do Universo possui um valor de $T_U = (1,42 \pm 0,04) \times 10^{10}$ anos, isto é, um valor de aproximadamente 14,2 bilhões de anos.

Gráfico 21: Análise das Curvas de Luz do RX J1131-1231



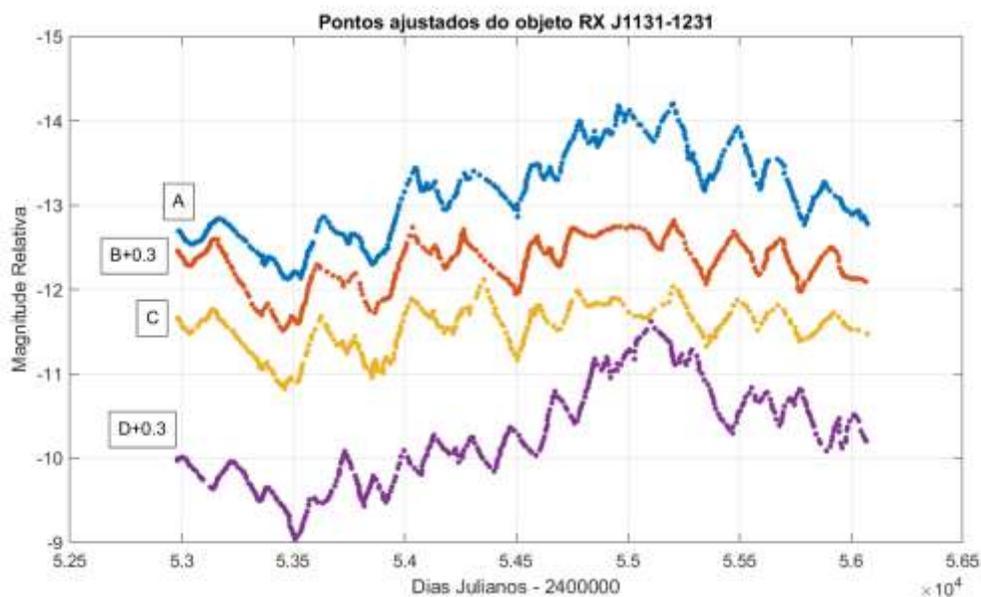
Fonte: Acervo do Autor

Tabela 1: Comparação entre as componentes do RX J1131-1231

	Δt (dias)	H_0 (Km/s/Mpc) $\pm 3,00$	T (s) $\pm 0,07$	T (anos) $\pm 0,02$
A e B	2,27	73,15	4,10E+17	1,30E+10
A e C	8,03	69,34	4,33E+17	1,37E+10
A e D	22,93	65,73	4,56E+17	1,45E+10
B e C	10,31	68,49	4,38E+17	1,39E+10
B e D	20,65	66,78	4,49E+17	1,42E+10
C e D	30,96	60,63	4,95E+17	1,57E+10

Já utilizando o segundo método, que é pelo algoritmo feito com base em redes neurais recorrentes, com aprendizado supervisionado, o Gráfico 21 se trata do ajuste destes pontos.

Gráfico 22: Pontos criados pela RNR do objeto RX J1131-1231



Fonte: Acervo do Autor

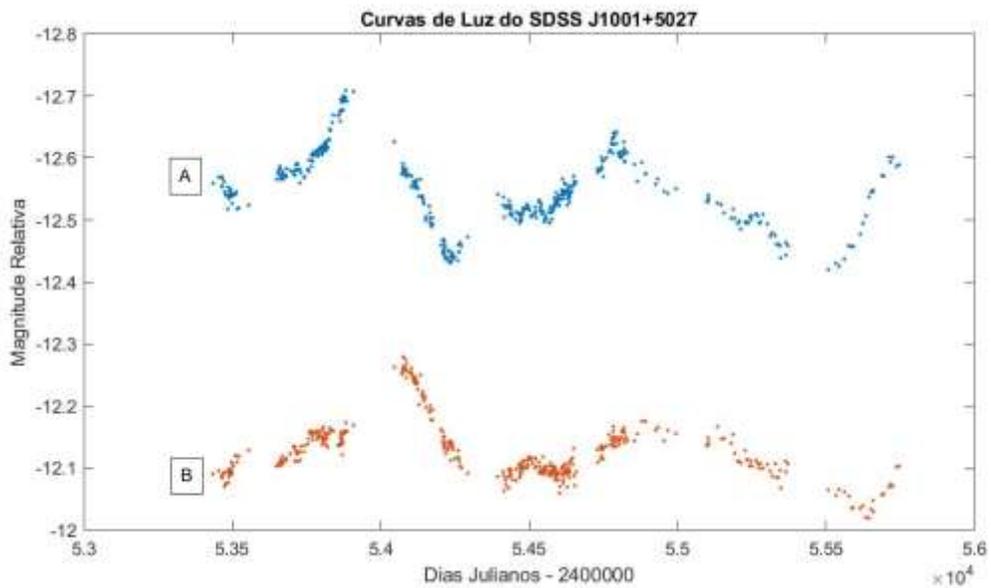
A partir do Gráfico 22, foi realizado a análise deste ajuste feito pela RNN e foi obtido um *time delay*, em média, para este candidato à lente gravitacional (RX J1131-1231) de aproximadamente $(14,82 \pm 0,93)$ dias.

Portanto, a constante de Hubble calculada foi de $H_0 = (69,4 \pm 2,7)$ Km/s/Mpc. Consequentemente, a idade do universo obtida foi de $T_U = (1,41 \pm 0,01) \times 10^{10}$ anos, isto é, 14,1 bilhões de anos com uma incerteza associada de 0,01 bilhão para mais ou para menos.

6.4 SDSS J1001+5027

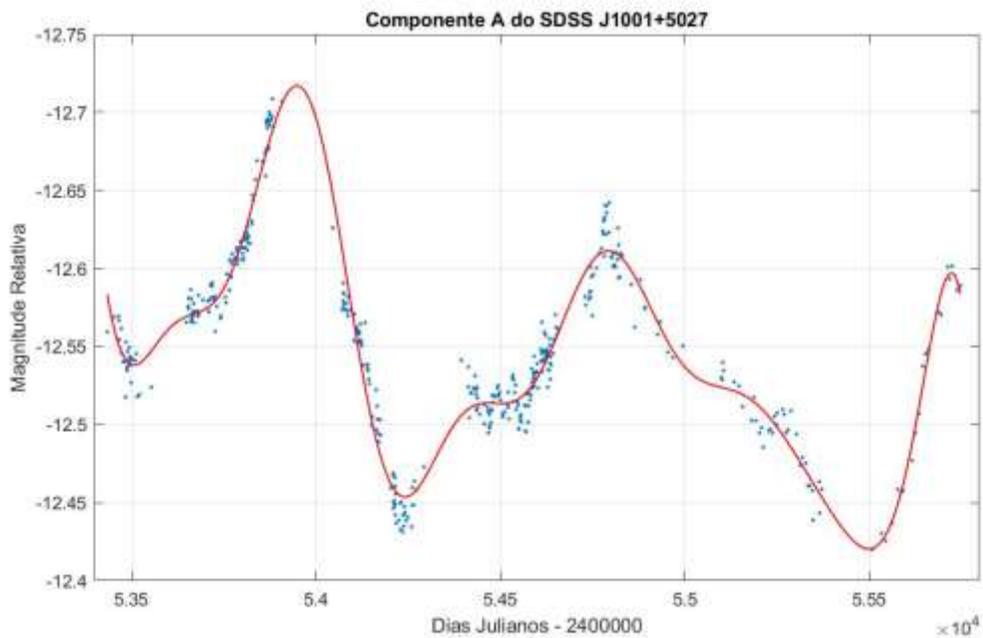
O gráfico obtido a partir dos dados do SDSS J1001+5027 pode ser observado no Gráfico 23.

Gráfico 23: SDSS J1001+5027



Fonte: Acervo do Autor

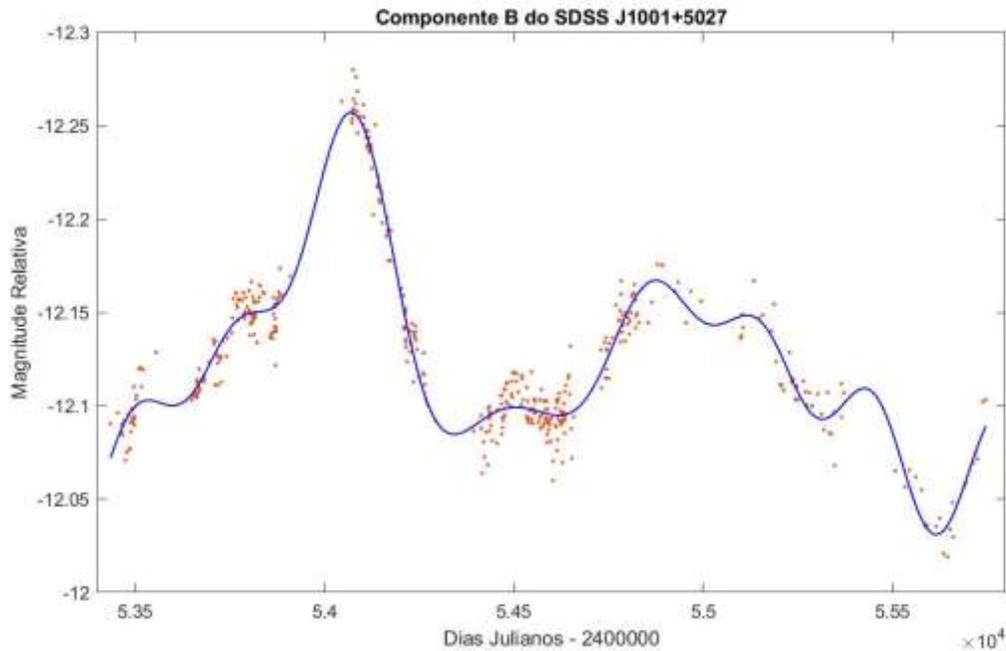
Gráfico 24: Componente A do SDSS J1001+5027



Fonte: Acervo do Autor

Para o ajuste ao conjunto de dados da componente A do SDSS J1001+5027 calculou-se o valor do R^2 de 0,9554.

Gráfico 25: Componente B do SDSS J1001+5027



Fonte: Acervo do Autor

Para o polinômio ajustado aos dados da Componente B do quasar-lente o valor do R^2 (coeficiente quadrático de regressão) foi calculado com o valor de 0,9529.

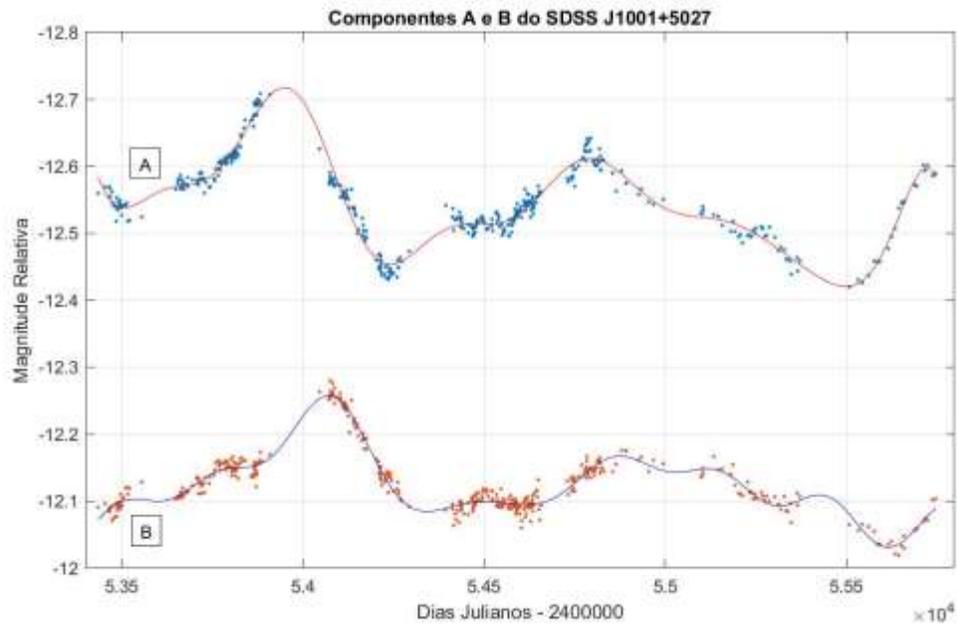
Como ambos os coeficientes quadráticos de regressão apresentam um valor relativamente alto, pode-se fazer a comparação entre a componente A e B deste quasar-lente, e calcular a defasagem de tempo entre o pico de máximo.

No Gráfico 26 é plotada as duas curvas de luz, e é perceptível uma grande diferença de tempo entre os picos de máxima magnitude.

Pode-se notar que há uma grande defasagem entre as curvas A e B do quasar-lente SDSS J1001+5027. Da mesma maneira que foi calculado esta defasagem nas lentes gravitacionais anteriores será feito neste. Logo, utilizando-se o polinômio ajustado pelo algoritmo calculou-se que a diferença de tempo entre os picos da componente A e B deste objeto é de cerca de $(122,52 \pm 3,59)$ dias.

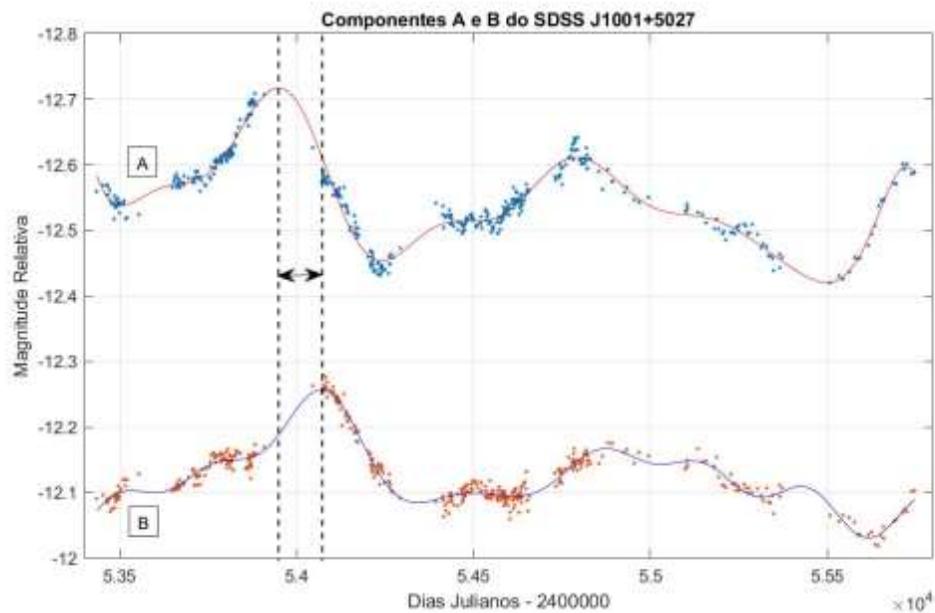
O “redshift” do objeto que está sob o efeito da lente gravitacional é de 1,838 (S. Rathna Kumar et al.,2013).

Gráfico 26: Análise das Curvas de Luz do SDSS J1001+5027



Fonte: Acervo do Autor

Gráfico 27: Análise das Curvas de Luz do SDSS J1001+5027



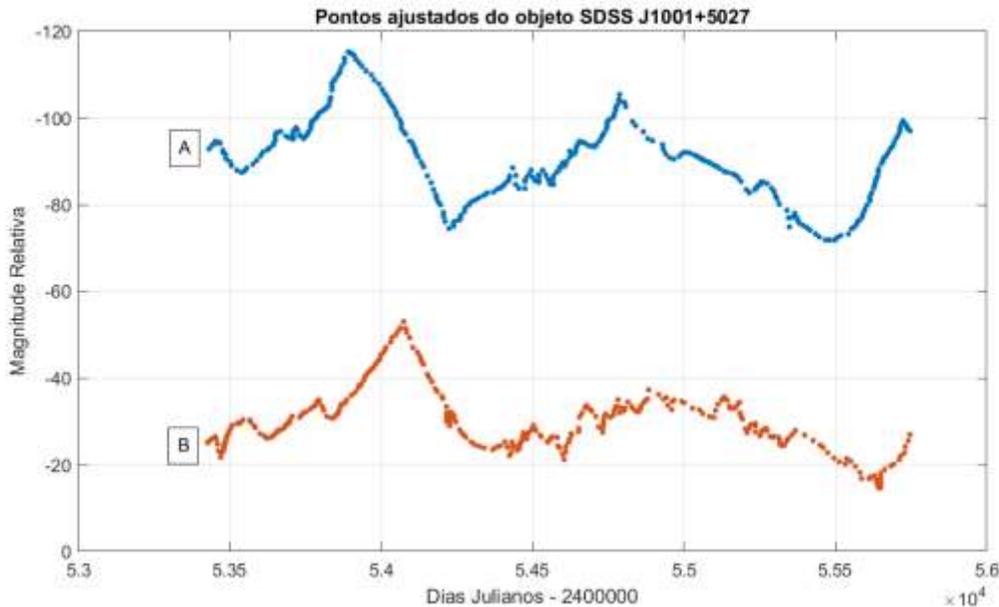
Fonte: Acervo do Autor

Utilizando-se a Equação 15 para determinar a constante de Hubble através dos dados obtidos pelo ajuste, pode-se calcular o valor de $H_0 = (75,0 \pm 5,4)$ Km/s/Mpc.

Consequentemente, através da Equação 16 pode-se estimar a idade do Universo com o valor de $T_U = (4,00 \pm 0,05) \times 10^{17}$ segundos, isto é, $T_U = (1,27 \pm 0,03) \times 10^{10}$ anos, ou seja, 12,7 bilhões de anos.

Já utilizando o segundo método, que é pelo algoritmo feito com base em redes neurais recorrentes, com aprendizado supervisionado, o Gráfico 28 se trata do ajuste destes pontos.

Gráfico 28: Pontos criados pela RNR do objeto SDSS J1001+5027



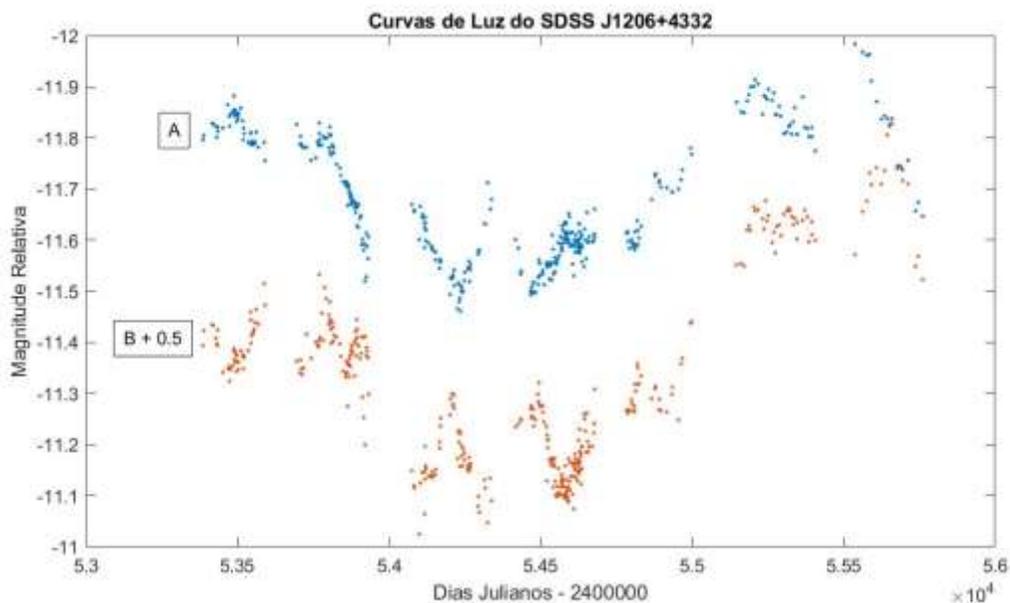
Fonte: Acervo do Autor

O ajuste feito no Gráfico 28 para o objeto SDSS J1001+5027 no intervalo de dados de aproximadamente 10 anos resultou que o *time delay* entre a componente A e B do quasar-lente é de aproximadamente $(142,09 \pm 1,78)$ dias. Consequentemente, o valor para a constante de Hubble obtido foi de $H_0 = (71,3 \pm 3,1)$ Km/s/Mpc.

Portanto, em mãos o valor calculado para a constante de Hubble, pode-se estimar a idade do Universo utilizando a Equação 14, logo $T_U = (4,33 \pm 0,01) \times 10^{17}$ segundos, isto é, $T_U = (1,37 \pm 0,01) \times 10^{10}$ anos, ou seja, 13,7 bilhões de anos, com uma incerteza associada de 0,01 bilhão para mais ou para menos.

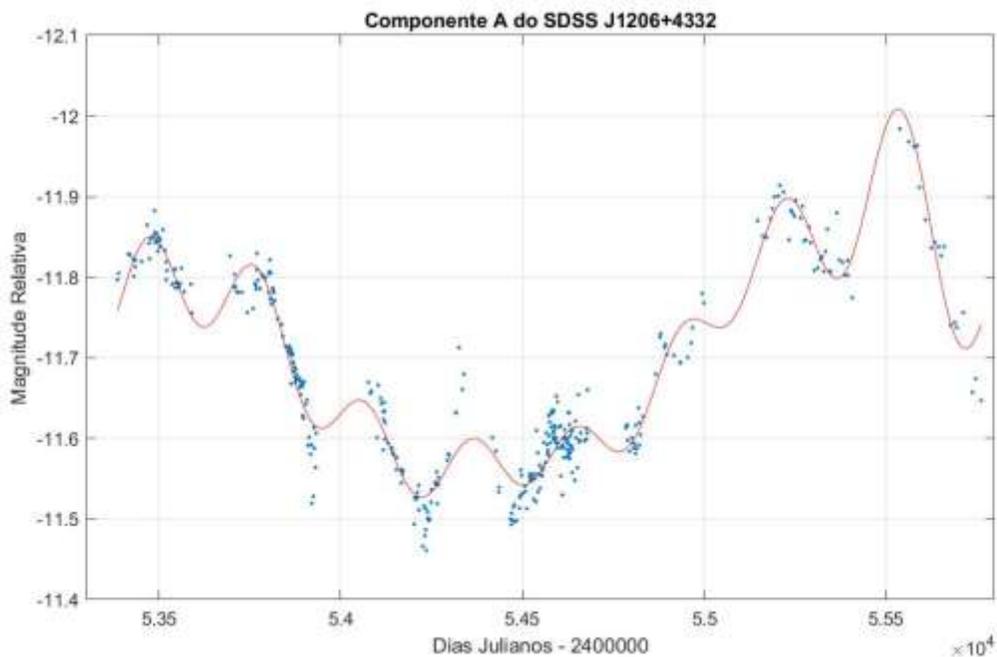
6.5 SDSS J1206+4332

Gráfico 29: SDSS J1206+4332



Fonte: Acervo do Autor

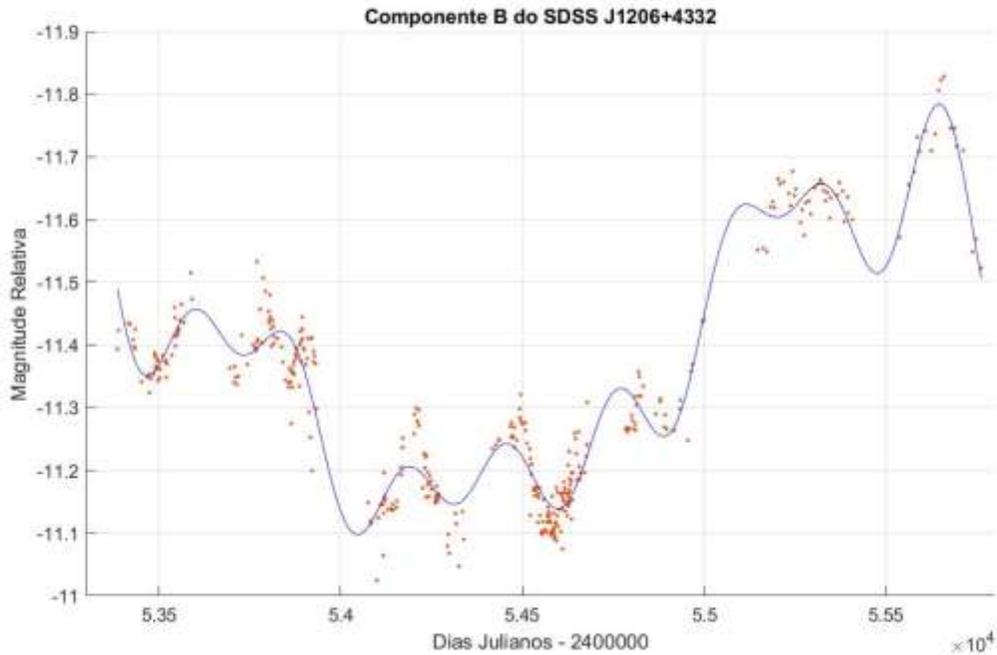
Gráfico 30: Componente A do SDSS J1206+4332



Fonte: Acervo do Autor

O valor do R^2 calculado pelo ajuste determinado no Gráfico 30 é de 0,9442.

Gráfico 31: Componente B do SDSS J1206+4332



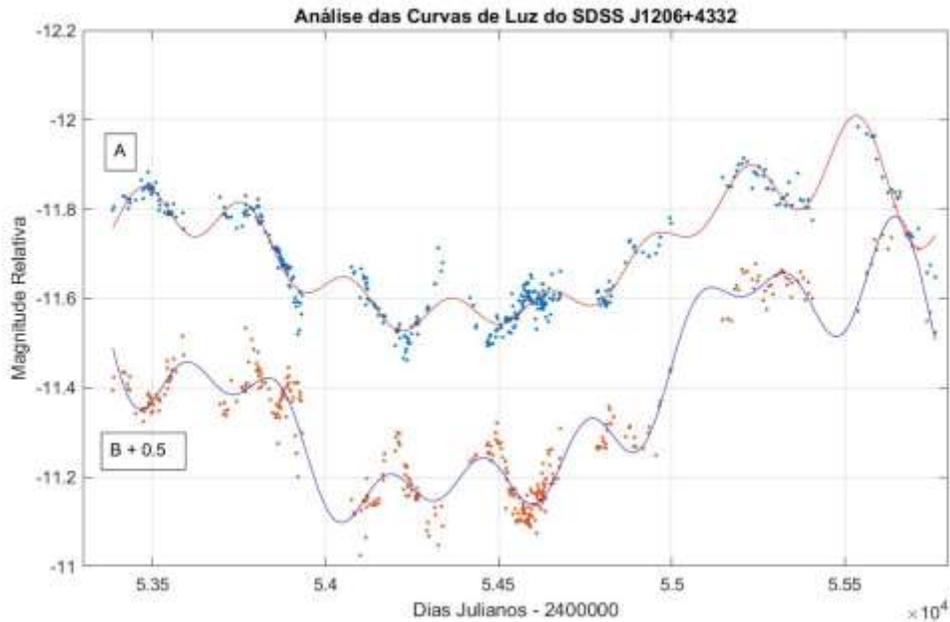
Fonte: Acervo do Autor

Para o ajuste do Gráfico 31 o Coeficiente Quadrático de Regressão (R^2) foi calculado com o valor de 0,9446. Consequentemente, pode-se dizer que ambos ajustes estão razoáveis para os conjuntos de dados, logo, uma comparação entre as curvas pode ser realizada.

Pode-se observar uma defasagem no valor de máxima magnitude relativa entre as componentes A e B (Figura 32).

Utilizando-se o algoritmo e o polinômio que o mesmo ajustou para as curvas A e B, pode-se estimar a diferença temporal entre o valor de máxima magnitude da componente A em relação à componente B, e este valor foi de $(121,62 \pm 4,32)$ dias. Consequentemente, utilizando-se a Equação 1 pode-se estimar a constante de Hubble com o valor de $H_0 = (73,5 \pm 5,7)$ Km/s/Mpc, para um $z = 1,789$ (Fonte: G.L.Q.D).

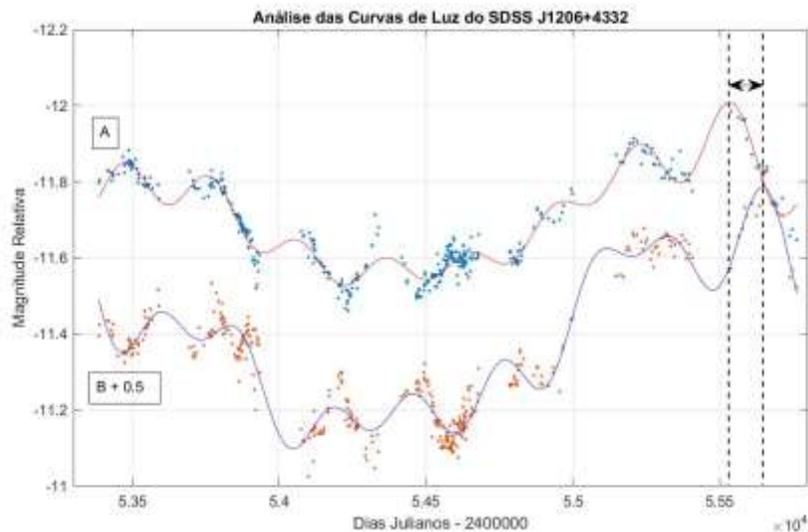
Gráfico 32: Análise das Curvas de Luz do SDSS J1206+4332



Fonte: Acervo do Autor

Estimando a idade do Universo através da Equação 2, obtêm-se que $T_U = (4,08 \pm 0,03) \times 10^{17}$ segundos, isto é, $T_U = (1,29 \pm 0,02) \times 10^{10}$ anos. Ou seja, pelo método que foi utilizado e os dados pesquisados na literatura foi calculado a idade do Universo de 12,9 bilhões de anos.

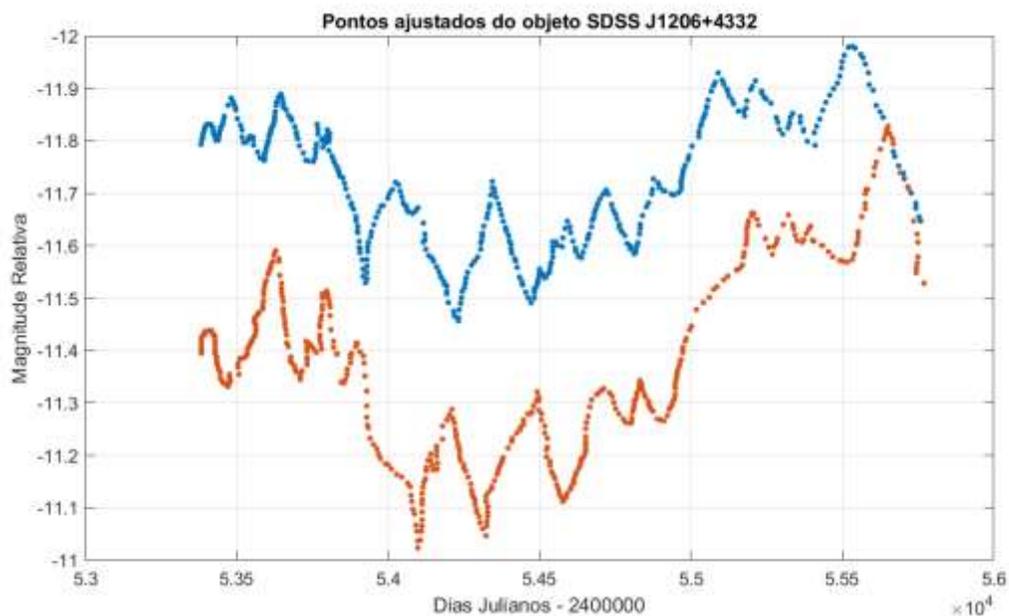
Gráfico 33: Análise das Curvas de Luz do SDSSJ1206+4332



Fonte: Acervo do Autor

Já utilizando o segundo método, que é pelo algoritmo feito com base em redes neurais recorrentes, com aprendizado supervisionado, o Gráfico 34 se trata do ajuste destes pontos.

Gráfico 34: Pontos criados pela RNR do objeto SDSS J1206+4332



Fonte: Acervo do Autor

A partir do Gráfico 34, que foi obtido utilizando a RNN, pode-se obter um valor para o *time delay* entre a componente A e B do quasar-lente SDSS J1206+4332 de aproximadamente $(134,62 \pm 1,19)$ dias.

Portanto, o valor da constante de Hubble para o *time delay* obtido é de $H_0 = (71,2 \pm 2,3)$ Km/s/Mpc. Consequentemente, a idade do Universo obtida em segundos foi de $T_U = (4,33 \pm 0,01) \times 10^{17}$ segundos, ou seja, $T_U = (1,37 \pm 0,01) \times 10^{10}$ anos.

Logo, a idade do Universo calculada utilizando o *time delay* da fonte SDSS J1206+4332 foi de 13,7 bilhões de anos, com uma incerteza para mais ou para menos de 0,01 bilhão.

7.0 RESULTADOS

Comparando-se os valores obtidos para as lentes gravitacionais e para a Constante de Hubble e conseqüentemente a Idade do Universo pode-se observar a Tabela 2. Para a lente gravitacional RX J1131-1231 foi utilizado o seu valor médio na Tabela 2, os valores para cada curva de luz da Constante de Hubble podem ser consultados na Tabela 1.

Tabela 2: Resultados para Fourier

Quasar-Lente	Constante de Hubble (Km/s/Mpc)	Idade do Universo (Bilhões de Anos)
HE0435-1223	$71,6 \pm 4,7$	$13,6 \pm 0,2$
HS 2209+1914	$62,8 \pm 4,8$	$15,5 \pm 0,3$
RX J1131-1231	$67,4 \pm 4,2$	$14,2 \pm 0,4$
SDSS J1001+5027	$75,0 \pm 5,4$	$12,7 \pm 0,3$
SDSS J1206+4332	$73,5 \pm 5,7$	$12,9 \pm 0,2$

Pode-se agora calcular a média dos valores obtidos para cada fonte a fim de determinar a constante de Hubble e sua incerteza associada.

Fazendo isto, pode-se obter que o valor da Constante de Hubble (H_0) determinado neste projeto é de $H_0 = (70,0 \pm 6,2)$ Km/s/Mpc. Conseqüentemente, a Idade do Universo (T_U) é de $T_U = (13,8 \pm 1,4)$ bilhões de anos.

O intervalo de confiança da média de H_0 e da idade do Universo foram estimados a partir do teste t-student para um nível de significância $\alpha = 5\%$ e para quatro graus de liberdade, uma vez que não é conhecida o valor da variância populacional de cada medida e teve-se que estimar a partir dos dados amostrais.

Já a Tabela 3 representa os resultados obtidos para a Constante de Hubble e a Idade do Universo utilizando a análise pelas Redes Neurais Recorrentes.

Tabela 3: Resultados para RRN

Quasar-Lente	Constante de Hubble (Km/s/Mpc)	Idade do Universo (Bilhões de Anos)
HE0435-1223	$69,4 \pm 1,7$	$13,6 \pm 0,1$
HS 2209+1914	$68,3 \pm 2,4$	$14,2 \pm 0,5$
RX J1131-1231	$69,4 \pm 2,7$	$14,1 \pm 0,1$
SDSS J1001+5027	$71,3 \pm 3,1$	$13,7 \pm 0,1$
SDSS J1206+4332	$71,2 \pm 2,3$	$13,7 \pm 0,1$

A partir da Tabela 3, pode-se calcular a média para a constante de Hubble e a idade do Universo de $H_0 = (69,92 \pm 5,55)$ Km/s/Mpc e $T_U = (13,91 \pm 0,07)$ bilhões de anos.

8.0 CONCLUSÕES

Foi possível estimar os atrasos temporais entre as componentes dos quasares candidatos à lentes gravitacionais utilizando-se a técnica de realizar um ajuste polinomial através da série de Fourier para cada curva de luz, determinando-se assim uma função harmônica para cada componente de cada quasar e com a mesma pode-se calcular no intervalo de dados que foram obtidos através do projeto COSMOGRAIL os pontos de máximos e mínimos das curvas e conseqüentemente determinar a defasagem no tempo de cada um destes pontos em relação a cada componente dos quasares.

Conseqüentemente, a partir do conhecimento de cada atraso temporal para cada quasar pode-se estimar a constante de Hubble sabendo-se que a mesma é diretamente proporcional à velocidade de dispersão da galáxia-lente e inversamente proporcional a defasagem no tempo em relação às componentes da fonte.

Obteve-se valores para a constante de Hubble para cada quasar candidato a lente gravitacional e finalmente, calculado a média entre cada fonte obtendo-se o valor de $H_0 = (70,0 \pm 6,2)$ Km/s/Mpc. Comparando-se o valor que foi obtido com o valor da constante de Hubble calculada por Kirshner, R em 2004 de $H_0 = (72 \pm 7)$ Km/s/Mpc podemos verificar que o valor que foi calculado no presente projeto possui um erro percentual de aproximadamente 2,7% em relação ao que foi estimado em 2004. Porém, devido à incerteza que se obteve de 6,2 Km/s/Mpc, este valor que foi calculado está dentro da margem de erro do valor estimado por Kirshner em 2004.

A partir do conhecimento da constante de Hubble calculada pode-se estimar a Idade do Universo, uma vez que esta constante possui dimensão de tempo. Realizando-se as conversões adequadas, pode-se estimar a idade do Universo de $(13,8 \pm 1,4)$ bilhões de anos. Comparando o valor que foi obtido com o calculado por Kirshner, R em 2004 de aproximadamente 13,2 bilhões de anos, alcançamos o valor esperado com um erro percentual de 4,4%.

Em suma, pode-se dizer que o projeto pelo pouco tempo de vigência que tomou obteve-se ótimos resultados para a Constante de Hubble e a Idade do Universo, porém grandes incertezas associadas devido ao método da série harmônica de Fourier não ser totalmente precisa, como pode-se observar os coeficientes quadráticos de regressão apresentam erros que não são nulos. O algoritmo LSTM que inicialmente foi proposto poderia otimizar o valor do R^2 para um valor muito próximo de 1.

Em relação aos valores obtidos pela RNN, pode-se observar que a precisão aumentou significativamente, uma vez que esta detecta os padrões da curva de luz e a completa. Desta forma, foi possível obter a constante de Hubble com um valor de $(69,92 \pm 5,55)$ Km/s/Mpc e conseqüentemente uma idade do Universo de 13,91 bilhões de anos com uma incerteza associada de 0,07 bilhão tanto para mais quanto para menos.

9.0 REFERÊNCIAS

ABBOTT, B. P. et al. OBSERVATION OF GRAVITATIONAL WAVES FROM A BINARY BLACK HOLE MERGER. (ligo Scientific Collaboration And Virgo Collaboration): Physical Review Letters, 2016.

BOVIN, V.; HOLICOW V. New COSMOGRAIL time delays of HE0435-1223: H0 to 3.8% precision from strong lensing in a flat Λ CDM model. Disponível em: <https://arxiv.org/abs/1607.01790>. Acesso em: 04 mar. 2020.

Chandra & XMM-Newton Provide Direct Measurement of Distant Black Hole's Spin".

Chandra X-ray Center. March 5, 2014. Retrieved March 5, 2014. Acesso em: 01 mar. 2020.

CHWOLSON, O. Über eine mögliche Form fiktiven Doppelsterne, Astr. Nachrichten, 221,329, 1924.

COOKE, J. H.; KANTOWSKI, R. Time delays for multiply imaged quasars, The Astrophysical Journal, 195, L11-L14, 1975.

COPPIN, B. Inteligência Artificial. Rio de Janeiro: Ltc, 2015. 636 p.

DINIZ, S. I. F. Lentes Gravitacionais: O Anel de Einstein PKS 1830-211. 2009. 76 f. TCC (Graduação) - Curso de Física, Centro de Ciências e Humanidades, Universidade Presbiteriana Mackenzie, São Paulo, 2009.

EDDINGTON, A. S., Space, time and gravitation, Cambridge University Press, Cambridge, 1920b.

EDDINGTON, A. S; M.A; M.Sc; F.R.S.Report on the relativity theory of gravitation.The physical society of London,1920a.

EINSTEIN, A. "Lens-like action of a star by the deviation of light in gravitational field", Science, 84, 506, 1936.

EINSTEIN, A. "Über den Einflub der Schwerkraft auf die Ausbreitung des Lichtes" Annalen der Physick, 35, 898, 1911.

EULAERS, E. COSMOGRAIL: the Cosmological Monitoring of Gravitational Lenses. XII. Time delays of the doubly lensed quasars SDSS J1206+4332 and HS 2209+1914. 2013. Disponível em: <https://ui.adsabs.harvard.edu/abs/2013A%26A...553A.121E/abstract>. Acesso em: 01 abr. 2020.

FAUSETT, L. Fundamentals of Neural Network. Florida Institute Of Technology: Prentice-hall, 1994.

FLORENTEN-NIELSEN, R. Determination of difference in light travel time for QSO0957-561 A, B. *Astr. Ap.*, 138, L19, 1984.

FRIAÇA, A. C. S., DAL PINO, E., SODRÉ JR., L., JATENCO-PEREIRA, V. (orgs). *Astronomia: uma visão geral do universo*. São Paulo: EDUSP, 2000.

HELERBROCK, R. "O que é efeito Doppler?"; *Brasil Escola*. Disponível em: <https://brasilecola.uol.com.br/o-que-e/fisica/o-que-e-efeito-doppler.htm>. Acesso em 13 de maio de 2020.

JACOBS, A. R. Increased rates of convergence through learning rate adaptation. Elsevier Ltd. University Of Massachusetts, p. 295-207. ago. 1988. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/0893608088900032>. Acesso em: 08 maio 2020.

KIRSHNER, R. P., AGUILERA, C., BARRIS, B., BECKER, A., CHALLIS, P., CHORNOCH, R., CLOCCHIATTI, A., COVARRUBIAS, R., FILIPPENKO, A. V., GARNAVICH, P.M., *et al.* (2003) *Bull. Am. Astron. Soc.* 202, 2308 (abstr.).

KLIMOV, Y. G. The deflection of light rays in the gravitational fields of galaxies, *Sov. Phys. Doklady*, 8, 119, 1963.

KUMAR, S. R. COSMOGRAIL: the Cosmological Monitoring of Gravitational Lenses XIV. Time delay of the doubly lensed quasar SDSS J1001+5027. Disponível em: <https://arxiv.org/abs/1306.5105>. Acesso em: 03 abr. 2020.

REDD, N. T. Monster Black Hole Spins at Half the Speed of Light. 2014. Disponível em: <https://www.space.com/24936-supermassive-black-hole-spin-quasar.html>. Acesso em: 17 mar. 2020.

REFSDAL, S. On the possibility of determining Hubble's parameter and the masses of galaxies from the gravitational lens effect, *MNRAS*, 128, 307, 1964b.

REFSDAL, S. The gravitational lens effect, *MNRAS*, 128, 295, 1964a.

SCHNEIDER, P., EHLERS, J., FALCO, E. E. GRAVITACIONAL LENSES. Editora Springer, 1992.

TIKHOV, G.A. Sur la deviation des rayons lumineux dans le champ de gravitation des étoiles. *Dokl. Akad. Nauk S.S.R.*, 16, 199-204, 1937.

WALSH, D.; CARSWELL, R. F.; WAYMANN, R. J. 0957 + 561 A, B – Twin quasistellar objects or gravitational lens *Nature*, v. 279, p. 381 – 384, 1979.

WISOTZKI, L. HE 0435-1223: a wide separation quadruple QSO and gravitational lens. Disponível em: <https://arxiv.org/abs/astro-ph/0207062>. Acesso em: 01 abr. 2020.

YOUNG, P. et al. The triple quasar Q1114+080A, B, C: A quintuple gravitational lens image? *The Astrophysical Journal*, v. 224, p. 723-735, 1981.

ZWICKY, F. "Nebulae as gravitational Lenses" *Phys. Rev.* 51, 290, 1937

ZWICKY, F. "On the possibility on detecting nebulae which act as gravitational lenses", *Phys. Rev.* 51, 679, 1937.